

CCAAATCGCCCGGTGCGGTGGTGCAGGGTCTCGGGCTAGTCAATGCGCTCCCGTCTCGGAGAC
TGCAGACTAAACCACTATTACTTGTTCAGAGCGTTCTGCTAATCTACACTTTTATTTTC
TGGATCACTGGCGTTATCCTTCTTGCACTTGGCATTTGGGGCAAGGTGAGCCTGGAGAATTA
CTTTTCTCTTTTAAATGAGAAGGCCACCAATGTCCCCTTCGTGCTCATTGCTACTGGTACCG
TCATTATTCTTTTGGGACCTTTGGTTGTTTTGCTACCTGCCGAGCTTCTGCATGGATGCTA
AAACTGTATGCAATGTTTCTGACTCTCGTTTTTTGGTFCGAAGTGGTCGCTGCCATCGTAGG
ATTTGTTTTTCAGACATGAGATTAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGC
AGTATAACTCTACAGGAGATTATAGAAGCCATGCAGTAGACAAGATCCAAAATACGTTGCAT
TGTTGTGGTGTACCGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGATT
TCCTAAGAGTTGCTGTAAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACA
ATGAAGGTTGTTTTATAAAGGTGATGACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGA
ATTTCTTTTGAGTGTCTTGCTTCCAAGTATTGGAATCTTTCTCGCCTACTGCCWCTCTCG
TGCCATAACAAATAACCAGTATGAGATAGTGTAAACCCAATGTATCTGTGGGCCTATTCCTCT
CTACCTTTAAGGACATTTAGGGTCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAAGTC
ACAAACCACTTACTGATGACCAAAAAACACCAAGTAGGTTGATTCGAATCAAGATGTAT
GTAGACCTAAACTACACCAATAGGCTGATTCATCAAGATCCCGTGCTCGCAGTGGGCTGAT
TCAATCAAGATGATGTTTGTATGTTCTTAAGTCCACCTTCATCCCATCATGTTAGATCG
TTGAAACCCGTGATCCCTCTGAACACTGGAAGAGCTAGTAAATTGTAATCAAGT

FIGURE 2

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA19902
><subunit 1 of 1, 245 aa, 1 stop, 1 unknown
><MW: -1, pI: 8.36, NX(S/T): 1
MASPSRRLQTKPVITCFKSVLLIYTFIFWITGVILLAVGIWGVSLNENYFSLLEKATNVPF
VLIATGTVIIILLGTFGCFATCRASAWMLKLYAMFLTLVFLVELVAAIVGFVFRHEIKNSFKN
NYEKALKQYNSTGDYRSHAVDKIQNTLHCCGVTDYRDWTDNTNYSEKGFPSCKCKLEDCTPQ
RDADKVNNEGCFIKVMTIIESEMGVVAGISFGVACFQLIGIFLAYCXSRITNNQYEIV
```

Important features of the protein:**Signal peptide:**

amino acids 1-42

Transmembrane domains:

amino acids 19-42, 61-83, 92-114, 209-230,

N-glycosylation site.

amino acids 134-138

Tyrosine kinase phosphorylation site.

amino acids 160-168, 160-169

N-myristoylation site.

amino acids 75-81, 78-84, 210-216, 214-220, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 69-80, 211-222

FIGURE 3

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACAAA
 GGAGTCGCGCCGCCGCCGCCGCCCTCCCTCCGGTGGGCCCGGGAGGTAGAGAAAGTCAGT
 GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT
 TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG
 GCCGGGGTAGGCTCTGGAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA
 GCCGAGAGGTTTTCCACCGAGGCCCGCGCTTGAGGGATCTGAAGAGGTTCTTAGAAGAGGGT
 GTTCCCTCTTTGGGGGTCTCACCAGAAGAGGTTCTTGGGGGTGCGCCCTCTGAGGAGGCT
 GCGGCTAACAGGGCCAGAATGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG
 GGAATAAGCTCTGCAACTTTCTTTGGCATTCAAGTTGTTAAAAACAAATAGGATGCAAATTCC
 TCAACTCCAGGTTATGAAAACAGTACTTGAAAACTGAAAACTACCTAAATGATCGTCTTTG
 GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGCCAGGGTCTGTTGTTGACTCTCGAAGAG
 CACATAGCCCACTTCTTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATCCTGTATCTG
 CCGAGATGACAGTGGAACAGATGACAGTGTGACACCCAACAGCAACAGGCCGAGAACAGTG
 CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCGTTCGGCCACCAAGGAGGGGC
 CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTTGGACACACT
 GGCAGTAATACGGACTCTTGTAGATAAGTAAAGTATCTGACTCACGGTCACCTCCAGTGGAAT
 GAAAAGTGTCTGCCCCGAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG
 CCAAGCCTTGTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAAA
 TGATAAGATTTGATGTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC
 TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

FIGURE 4

MIVFGWAVFLASRSLGQGLLLTLEEHIHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ
AENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKQNVLDGLVLDTLAVIRTLVDKO

Signal peptide:

amino acids 1-16

Casein kinase II phosphorylation site.

amino acids 22-26, 50-54, 113-117

N-myristoylation site.

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

FIGURE 5

GGCACGAGGCGCTGTCCACCCGGGGCGTGGGAGTGAGGTACCAGATTTCAGCCCATTGCGC
 CCGACGCCTCTGTTCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCTTAACGGACTG
 CAAG**ATG**GAGGAAGGCGGGAACCTAGGAGGCTGATTAAGATGGTCCATCTACTGGTCTTGT
 CAGGTGCCTGGGGCATGCAAATGTGGGTGACCTTCGTCTCAGGCTTCCTGCTTTTCCGAAGC
 CTTCCCCGACATACCTTCGGACTAGTGCAGAGCAAACCTTCCCTTCTACTTCCACATCTC
 CATGGGCTGTGCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCA
 CATTCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCCTGAGCCTTACGCTGGCCACTGTCAAC
 GCCCCTGGCTGGAACCCCGCACCACAGCTGCCATGTGGGCCCTGCAAACCGTGGAGAAGGA
 GCGAGGCCTGGGTGGGGAGGTACCAGGCAGCCACCAGGGTCCCGATCCCTACCGCCAGCTGC
 GAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCTTCCGCTACCATGGGCTGTCC
 TCTCTTTGCAATCTGGGCTGCGTCCTGAGCAATGGGCTCTGTCTCGCTGGCCTTGCCCTGGA
 AATAAGGAGCCTC**TAG**CATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAA
 AAAAAAAAAA

FIGURE 6

```

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107
<subunit 1 of 1, 231 aa, 1 stop
<NX(S/T): 0
MEEGGLGGLIKMVHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFFPYFHISM
GCAFINLCLILASQHAWAQLTFWEASQLYLLFLSLTLATVNARWLEPRTTAAMWALQTVEKER
GLGGEVPGSHQGPDYPYRQLREKDEPKYSALRQNFTRYHGLSSLNLGCVLSNGLCLAGLALEIRSL

```

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 86-103, 60-75

Casein kinase II phosphorylation site.

amino acids 82-86

Tyrosine kinase phosphorylation site.

amino acids 144-151

N-myristoylation site.

amino acids 4-10, 5-11, 47-53, 170-176, 176-182

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 54-65

G-protein coupled receptors proteins.

amino acids 44-85

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 7

AATTCAGATTTTAAGCCATTCTGCAGTGGAAATTTTCATGAAGTAGCAAGAGGACACCATCTT
 CTTGTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGAAAAATGCTCTTTTGGGT
 GCTAGGCCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAG
 ACATCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCC
 AGAATCTTTGATAAAAAGGATTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAAC
 AGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTAAGTGTGCTTCTGGATGTGACCGACCCAG
 AGAATGTCAAGAGGACTGCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGT
 CTGATCAATAATGCTGGTGTTCCCGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGA
 CTACAGAGAACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTC
 CTTTGGTCAAGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTGGAGGTCGCCTTGCA
 ATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAGGTTTCAATGACAGCTTAAG
 ACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACAGGATTGTTCAAAACAA
 ACTTGGCAGATCCAGTAAAGGTAATTGAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCA
 GACATCAAAACAATATGGAGAAGGTTACATTGAAAAAGTCTAGACAAACTGAAAGGCAA
 TAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCAGCTCTAACAA
 GTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAATTTTCTGGATACCTCTG
 TCTCACATGCCAGCAGCTTTCGAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAA
 TCCCAAGGCAGTGTGACTCAGCTAACCAAAATGTCTCCTCCAGGCTATGAAATTGGCCGAT
 TTCAAGAACACATCTCCTTTTCAACCCATTCTTATCTGCTCCAACCTGGACTCATTTAGA
 TCGTGCTTATTTGGATTGCAAAAGGAGTCCACCATCGCTGGTGGTATCCAGGGTCCCTG
 CTCAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCTGT
 ATTTAGGCTTTGCTGTCTGGTGTGATGTAAGGGAATTGAAAGACTTGGCCATTCAAATG
 ATCTTTACCGTGGCTGCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTT
 AAGTATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAA
 AAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406
><subunit 1 of 1, 319 aa, 1 stop
><MW: 35227, pI: 8.97, NX(S/T): 3
MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFTTGCDSGFGNLAARTFDKKGPHVIAACLT
ESGSTALKAETSERLRTVLLDVTDPENVKRTAQWVKQVGEKGLWGLINNAGVPGVLAPTDW
LTLEDYREPIEVNLFGLISVTNLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGF
NDSLRRDMKAFGVHVSCEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLD
KLKGNKSYVNMDLSPVVECDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQK
AELANPKAV

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 9

GCGGGCTGTTGACGGCGCTGCGC**ATG**GGCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTT
 CCTCTCAGTCGGACTTCTGTACGCGCCAGTGGGCGGGGCCCTTGGGGCGTCCGCCACCACT
 GTAGTCATGTACCCACGCGCGCGCCGCCCTCATCGGCACTTCATCTCGGTGACCGTGA
 CTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGGCGCGCTCGTGGGAGGAAATGGA
 AGCAACTGTGCGAGATTGCAGCGGAATATGATTCTCTTCTCCTTGCCTTTCTGCTTTCTGTT
 GGACTCCTCTTCTACATCAACTGGCTGACCATTGGAAAGCTTGGCTTTCAAGCTAGAGGA
 AGAGCAGAAGATGAGGCCAGAAATTGCTGGGTTAAACCAGCAAATCCACCCGCTCTTACCAG
 CTCCCTCAGAAGGCGGACACCGACCCGTGAGAACTTACCTGAGATTTTCGTACAGAAAGACACAA
 AGACACATCCAGCGGGGACCACTCACCTGCAGATTAGACCCCCAAGCCAAGACCTGAAGGA
 TGGGACCCAGGAGGAGGCCACAAAAAGGCAGAAGCCCTGTGGATCCCCGCCGGAAGGAG
 ATCCGCAGAGGACAGTCACTAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGGCACCGAG
 CTCCCTTCAAGAAGAGCAGAAGTGCCCAACCAAGCTCCCTGCCACCGGCCAGGACACAGGG
 CACACCAAGTGCATCTGAACATATCGCCAGAAAGGCGGTGATTGACGCTTCTCCTGCATGCATGGA
 AAGGATACCCGAAGTTTGCATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCCTTCAGT
 GAGTGGTTTGGCCTCGGTCTCACACTGATCGACGCGCTGGACACCATGTGGATCTTGGGTCT
 GAGGAAAGAATTGAGGAAGCCAGGAAGTGGGTGTGGAAGAAGTTACACTTTGAAAAGGACG
 TGGACGTCAACCTGTGTTGAGAGCAGCATCCGCATCCTGGGGGGCTCCTGAGTGCCTACCAC
 CTGTCTGGGGACAGCCTTCTCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCTGCT
 TCTCAGAACCATTCCAAGATTCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCACC
 CGCCACGTTGACCTCCGACAGCACTGTGGCCGAGGTGACCAGCATTCAGCTGGAGTTCCGG
 GAGCTCTCCCGTCTCACAGGGGATAAGAAGTTTCAAGAGGCAGTGGAAGAAGTGACACAGCA
 CATCCACGGGCTGTCTGGGAAGAAGGATGGGCTGGTGCCCATGTTCAATACCCAGCAAGT
 GCCTCTTACCCACCTGGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTAC
 CTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGA
 AGCCATCGAGGGTGTGAGAAGCAGCTGCTGCGGCATCCGAGGCCAGTAAGCTCAAGCTTTG
 TGGGGGAGCTTGCCACGCGCGCTTCAGTGCCAAGATGGACCACTGGTGTGCTTCTCGCCA
 GGGACGCTGGCTCTGGGCGTACCACGGCCTGCCCGCCAGCCACATGGAGCTGGCCAGGA
 GCTCATGGAGACTTGTTACCAGATGAACCGGCAGATGGAGACGGGGCTGAGTCCCGAGATCG
 TGCATTCAACCTTTACCCCCAGCCGGGCGCTCGGGACGTGGAGGTCAAGCCAGCAGACAGG
 CACAACCTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCTCACAGGGGA
 CCGCAAAATACAGGACTGGGGCTGGGAGATTCTGCAGAGCTTCAGCCGATTACACGGGTCC
 CCTCGGTGGCTATTCTTCCATCAACAATGTCCAGGATCCTCAGAAGCCCGAGCCTAGGGAC
 AAGATGGAGAGCTTCTTCTGGGGGAGACGCTCAAGTATCTGTTCTTGCTCTTCTCCGATGA
 CCCAAACCTGCTCAGCCTGGACGCTACGTGTTCAACACCGAAGCCCACTCTGCTTCTATCT
 GGACCCCTGCT**AGG**GTGGATGGCTGCTGGTGTGGGCACTTCGGGTGGCGAGGACCTTG
 CTGGGTCTGTGGCATTTCCTCAAGGGCCACGCTAGCACCGGCAACCGCAAGTGGCCAGGCT
 CTGAACCTGGCTCTGGGCTCCTCCTCGTCTCTGCTTTAATCAGGACACCGTGAGGACAAGTGA
 GGCCGTGAGTCTTGGGTGTGATGCGGGGTGGGCTGGGCGCTGGAGCCTTCGGCTTCTCCTC
 CAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAGGTCTCTGTGGGCGGACCA
 GAGGGGGCTTCGAGGTGGTCCCTGGTACTGGGTGACCGAGTGGACAGCCAGGGTGCAGC
 TCTGCCCGGGCTCGTGAAGCCTCAGATGTCCCAATCCAAGGGTCTGGAGGGGCTCCCGTGA
 CTCAGAGGCTGAGGCTCCAGGGCTGGCTCTGGTGTTCACAAAGCTGAGCTCAGGATCTCTC
 CTGGCCGCCCCGACGGGGCTTGGAGGCTGGACGGCAAGTCCGCTTAGCTCAGCGGCCCT
 CCAGTGGAAATGGGTCTTTTCGGTGGAGATAAAAGTTGATTGCTCTAACCGCAA

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529
><subunit 1 of 1, 699 aa, 1 stop
><MW: 79553, pI: 7.83, NX(S/T): 0
MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPPHRDFISVTLSFGESYDN
SKSWRRRSCWRKWKQLSRLQRNMILFLAFLFCGLLFYINLADHWKALAFRLEEEQKMRPE
IAGLKPANPFVLPAPQKADTDEENLFEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTEEAT
KRQEAPVDPRPEGDFQRTVISWRGAVIEPEQGTELPSSRAEVPTKPPLPPARTQGTPVHLNY
RQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTMWILGLRKEFEEA
RKWVSKKLHFEKDVDVNLFESTITILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKI
PYSDVNIGTGVAHPFRWTS DSTVAEVT SIQLEFRELSRLTGDKKFQEA VEKVTQH IHGLSGK
KDGLVPMFINTHSGLFTHLGVFTL GARADSYEYLLKQW IQGGKQETQLLEDYVEA IEGVRT
HLLRHSEPSKLT FVGELAHGRFS AKMDHLVCFLPGTLALGVYHGLPASHMELAEQELMETCYQ
MNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLRLRPETVESLFYLYRVTGDRKYQDWG
WEILQSFSRFRTRVPSGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDDPNLLSLD
AYVFNTEAHPLPIWTPA

Important features of the protein:

Transmembrane domain:

amino acids 21-40 and 84-105 (type II)

FIGURE 11

10156742401

GGCGCCGCTAGGCCCGGGAGGCCGGGCCGGGCTGCGAGCGCTGCCCCATGCGCCGC
 CGCCTCTCCGCACG**ATG**TTCCTTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGC
 AGGTCGGGTTGCTCTCCGGCGGCCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTCTGTGGC
 CTGCCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGG
 CCCGGGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCGTGCTGCCCCCAGAG
 CCGCCCCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCACCCTGGCAGTGCTGGT
 GCCCTTCCGCGAACGCTTCGAGGAGCTCCTGGTCTTCGTGCCCCACATGCGCCGCTTCTCTGA
 GCAGGAAGAAGATCCGGCACCATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAAC
 CGGGCAGCGCTCATCAACGTGGGCTTCTCTGGAGAGCAGCAACAGCACGGACTACATTGCCAT
 GCACGACGTTGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCCTGAGGCTGGGC
 CCTTCCACGTGGCCTCCCCGGAGCTCCACCTCTCTACCACTACAAGACCTATGTCGGCGGC
 ATCTTGCTGCTCTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGG
 CTGGGGCCGCGAGGACGACGAGTTCTACCGCGCATTAAGGGAGCTGGGCTCCAGCTTTTCC
 GCCCCCTCGGAATCACAACTGGGTACAAGACATTTGCCACCTGCATGACCCAGCGGTGGCGG
 AAGAGGGACCAGAAGCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGG
 AGGCCTGAACACTGTGAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCC
 CCTGCACTGTCTCAACATCATGTTGGACTGTGACAAGACCGCCACACCTTGGTGACATTC
 AGCT**TGA**GCTGGATGGACAGTGAGGAAGCCTGTACCTACAGCCATATTGCTCAGGCTCAGGA
 CAAGGCTCAGGTCGTGGGCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCA
 AGCTACGCAATTGCAGCCACCGGCCCAAGGCAGGCTTGGGCTGGGCCAGGACACGTGGG
 GTGCCTGGGACGCTGCTTGCCATGCACAGTGATCAGAGAGAGGCTGGGGTGTGCTCTGTCCG
 GGACCCCCCTGCCTTCTGTCTACCCCTACTCTGACCTCCTTCACGTGCCAGGCCGTGGG
 TAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCTACTCTGACCTCCTTCACGTGCC
 AGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCCCCAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 12

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531
><subunit 1 of 1, 327 aa, 1 stop
><MW: 37406, pI: 9.30, NX(S/T): 1
MFPSSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLFVACLSSLGFFSLLWLQLSCSGDVARAVR
GQGQETSGPPRACPEPPPPEHWEEDASWGPRLAVLVFPRFRFEELLVFVPHMRRFLSRKKI
RHHIYVLNQVDHFRFNRAALINVGFLSSNSTDYIAMHDVDLLPLNEELDYGFPEAGPFHVA
SPELHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDFYYRRIKGAGLQLFRPSGI
TTGYKTFRHLHDPAPWRKRDQKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVL
NIMLDCDKTATPWCTFS
```

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 29-49 (type II)

N-glycosylation site.

amino acids 154-158

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 27-31

Tyrosine kinase phosphorylation site.

amino acids 226-233

N-myristoylation site.

amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

FIGURE 13

CAATGTTTGCCTATCCACCTCCCCAAGCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCT
GCTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCT
CTTCTGCCACTGACGCCCCATCAGGGATTGGGCCTTCTTCCCCCTTCCTTTCTGTGTCTC
CTGCCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGT
GGGGGATGGCT**TAA**GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGG
GCTGCCTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGC
CCTGTCATTAAAGAAATGTTCCCTTCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGT
AAATATGTCTTTATAATAAACAGTTAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 14

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56862
```

```
<subunit 1 of 1, 73 aa, 1 stop
```

```
<MW: 7879, pI: 7.21, NX(S/T): 0
```

```
MLLLTLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQ  
PRGEGEKVGDG
```

Important features:**Signal peptide:**

```
amino acids 1-15
```

Growth factor and cytokines receptors family:

```
amino acids 3-18
```

FIGURE 15

GGGACCCATGCGGCCGTGACCCCGGGCTCCCTAGAGGCCAGCGCAGCCGAGCGGACAAAG
 GAGCATGTCCGGCCGCGGGAAGGCCCGTCCCTCGGCCGCCATAAGGCTCCGGTCGCCGCTGG
 GCCCGCCGCGCGCTCCCTGCCCGCCCGGGCTCCGGGGCGGCCGCTAGGCCAGTGCGCCGCCG
 CTCGCCCGGAGGCCCGCGGCCCGCAGCATGAGCGACCCGCGGCCGCGGGGCCGCGCGCA
 GCCGCCGCTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGAGGCGCGCGG
 CGGGCGGCCGCGCGGCGCTGCCCGCGGCTGCAAGCAGCATGGCGGCCCGGAGGGGCTGGC
 AGGCGCGCGGCCGCGGCCGAGGGCAAGGTGGTGTGACAGCGCTGGAACCTCGCGCAGGTCCCT
 GCCCCAGATACTGTGCCCAACCGCACGGTACCCTGATTCTGAGTAACAATAAGATATCCG
 AGCTGAAGAATGGCTCATTTCTGGGTTAAGTCTCCTTGAAAGATTGGACCTCCGAAACAAT
 CTTATTAGTAGTATAGATCCAGGTGCCCTTCTGGGACTGTCACTCTAAAAAGATTGGATCT
 GACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTGAGGACTCACC AATCTGGTTC
 GGCTAAACCTTTCCGGGGAATTTGTTTTCTTCATTATCTCAAGGAACCTTTGATTATCTTGGC
 TCATTACGGTCTTTGGAATTCAGACTGAGTATCTTTGTGTGACTGTAACTACTGTGGAT
 GCATCGCTGGGTAAGGAGAAGAACATCACGGTACGGGATACCAGGTGTGTTTATCCTAAGT
 CACTGCAGGCCCAACAGTACAGGCGTGAAGCAGGAGCTGTTGACATGCGACCCCTCCGCTT
 GAATGCGCGCTTTCTACATGACTCCATCTCATCGCCAAAGTTGTGTTTGAAGGAGACAGCCCT
 TCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTGTTGGTATCAGGATG
 GGAGAATAGTTGAACCGATGAATCGCAAGGTATTTTGTGTGAAAAGAACATGATTCACAACT
 TGTCCTTGATTGCAAGTGCCTTAACCATTTCTAATATTCAAGCTGGATGCTACTGGAAATTTG
 GGGCTGTCTATGTCAGACCAAACTGGGAATAATACGAGGACTGTGGATATTGTGGTATTAG
 AGAGTTCTGCACGATCTGTCTCCAGAGAGGGTGGTAAACAACAAGGTGACTTCAGATGG
 CCGAGAATTTGGCAGGCAATTAATGCATATCTGCAGTGTACCGCGCAACACCTGAGCATGG
 GATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGGCGCAGATGTGATAGAGGTGGCT
 TTTGGGCAGATGATGATTATTCTCGCTGTCAATGCAATGATGTACATAGATTCTTTAT
 ATGTTTAATCAGATGCCCTCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGC
 TTACACTGTGGAAGCAGCCAACCTTTCTGACAAAATGGATGTTATATTGTGGCAGAAATGA
 TTGAAAATTTGGAAGATTACCAAGGAGGAAAAATCAAAAGAGCTAGGTGACGTGATGGTT
 GACATTGCAAGTAACATCATGTTGGCTGATGAACGTGTCTGTGGCTGGCGCAGAGGGAAGC
 TAAAGCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGCGTAGCGGTG
 GAGCTCACGTTTATTCACATATTCACCAATATTGCTCTGGAAGCTTATGTCAATCAAGTCT
 ACTGGCTTCACGGGGATGACCTGTACCCTGTTCAGAAAGTGGCAGCCTCTGATCGTACAGG
 ACTTTCCGATTCTGGAGGCGGGATCCAGAGGGAACCTGGATAAGCAGCTGAGCTTTAAGT
 GCAATGTTTCAAATACATTTTCGAGTCTGGCCTAAAGGTATGTTACATTCTGCAATCATTT
 AAGACTATTTACAGTAAAATTAGAATGCTCCAAATGTTCTGCTTCGCAAAATAACCTTATTA
 AAGACTTTTCTTTTGGAGGAAGATAGGTATTGCTTTGCTAGTGTTTTAAAGAAAACCTA
 ACCAGGAAGAACTGCATTACGACTTTCAAGGGCCCTAGGCATTTTGGCTTTGATTCCCTTT
 CTTTCACATAAAAAATCAGAAATTACATTTTATACTGCAGTGGTATAAATGCAAAATATACT
 ATTGTTACATGTGAAAAAATTTTATTTGACTTAAAGGTTATTTATTGTTTTTGTCTCT
 GATTTTAAAGACAATAAGATGTTTTCATGGGCCCCTAAAGATATCATGAGCCTTTGGCACTGC
 GCCTGCCAAGCCTAGTGGAGAAAGTCAACCTGAGACCAAGGTGTTTAAATCAAGCAAGCTGTAT
 ATCAAAATTTTGGCAGAAAAACACAAATATGTCATATATCTTTTAAAAAAAGAAATTTCA
 TTGAAGCAAGCAAAATCAGAACTTTTACTGATTTTAAATTTGGTGGCTTTAGATATATT
 GACTACACTGTATTGAAGCAATAGAGGAGGCACAACCTCCAGCACCTTAATGGAACACATT
 TTTTCTCACTTAGCTTTCTGTGGGATGTGTAATTGTATTCTCTGCGGTTTTAATCTCACAG
 TACTTTATTTCTGTCTTGTCCCTCAATAATATCAAAACAATTTCCAGTCAATTTTAATGGC
 TGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTAGTGTGAGCAGCTCAATAAATA
 TTGAATGAATGAACGAAAAAAAAAAAAAAAA

FIGURE 16

MEPPGRRRRGRAQPPLLLPLSLALLALLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGK
 VVCSSLELAQVLPPDTPNRTVTILSNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGA
 FWGLSSSLKRLDLTNNRIGCLNADIFRGLTNLVRNLNSGNLFSSLSQGTFDYLASLSLEFQT
 EYLLCDCNIIWMHRWVKEKNITVRDTRCVYPKSLQAQPVTVGVKQELLTCDPPELPSFYMTP
 SHRQVVFEGDSLFPQCMASYIDQDMQVLWYQDGRIVETDESQGIFVEKNMIHNCSLIASALT
 ISNIQAGSTGNWGCHVQTKRGNNTRTVDIVVLESSAQYCPPERVVNNKGDFRWPRTLAGITA
 YLQCTRNTHGSGIYPGNPQDERKAWRRCDRGGFWADDDYSRCQYANDVTRVLYMFNQMPNL
 TNAVATARQLLAYTVEAANFSKMDVIFVAEMIEKFGRFTKEEKSKELGDVMVDIASNIMLA
 DERVLWLAQREAKACSRIVQCLQRIATYRLAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCT
 VFQKVAASDRTGLSDYGRRDPEGNLDKQLSFKCNVSNTFSSLALKVCYILQSFKTIYS

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 13-40 (type II)

N-glycosylation site.

amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336,
 433-437, 453-457, 592-596

N-myristoylation site.

amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57,
 57-63, 99-105, 123-129, 142-148, 162-168, 317-323, 320-326,
 384-390, 403-409, 554-560

FIGURE 17

GCCTGGGGATCTCTAGGAGCTCGAAGGTGGTGCTGGGCCTCTCGGTGCTGCTGACGGCGGCC
ACAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTAT
CAGAGACATTGAGAGGCAAATTCGAAAAAAGAAAAACATTCGTCTTTTGGGAGAACAGATTA
TTTTGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAAGGATCTCAAAAA
TCTGACTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTTGAT
GGAGAGTAGCTTAGTAGTATCTTCATCTTTTTTTTTGGTCACTGTCCTTTTAACTTGATCA
AATAAAGGACAGTGGGTATATAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGA
GTGTGGGCAGACACTTTTTTGGAAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGG
TCACTGTCCAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGA
TGGATAGTAGCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACA
TCGACCTCAGAAGGACTGGAGGAAGGTGAAGTGGAGGGAGAGACGCTCCTGATCGTCGAATCC

FIGURE 18

MSRSSKVVVLGLSVLLTAATVAGVHVKKQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILT
EQLEAEREKMLLAKGSQKS

Signal peptide:

amino acids 1-21

18/249

FIGURE 19

CTGTCGTCTTTGCTTCAGCCGAGTCGCCACTGGCTGCCTGAGGTGCTCTTACAGCCTGTTT
 CAAGTGTGGCTTAATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCTCTGCTAAGACC
 GCTGCC**ATG**CCAGTGACGGTAACCCGACCACCATCACAACCACCACGACGTATCTTCGGG
 CCTGGGGTCCCCATGATCGTGGGTCCCCTCGGGCCCTGACACAGCCCCGGGTCTCCTTCGC
 CTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGTGGGCGCCTGGAC
 GGGGTCCATGGGCAACTGGTCCATGTTACCTGGTGTCTTCTGCTTCTCCGTGACCTGATCA
 TCCTCATCGTGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGGCGCAACTTCCCC
 ATCACCTTCGCCTGCTATGCGGCCCTTCTTGCCTCTCGGCCCTCCATCATCTACCCACCAC
 CTATGTCCAGTTCTCTGTCCCACGGCCGTTTCGCGGGACCACGCCATCGCCGCCACCTTCTTCT
 CCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTGGACCCGGGGCCCGGCCGCGAG
 ATCACTGGCTATATGGCCACCGTACCCGGGCTGCTGAAGGTGCTGGAGACCTTCGTTGCCTG
 CATCATCTTCGCGTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGCCCTGGAGTGGT
 GCGGTGGCGGTGTACGCCATCTGCTTCATCTAGCGGCCATCGCCATCTGCTGAACCTGGGG
 GAGTGCACCAACGTGCTACCCATCCCTTCCCCAGCTTCTGTGCGGGCTGGCCTTGTCTGTC
 TGTCTCTCTATATGCCACCGCCCTTGTCTCTGGCCCCCTACCACTTCGATGAGAAGTATG
 GCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGAGCCATGCCTACTACGTGTGT
 GCCTGGGACCGCCGACTGGCTGTGGCCATCTGACGGCCATCAACCTACTGGCGTATGTGGC
 TGACCTGGTGCACCTCTGCCACCTGGTTTTTGTCAAGGTCT**TAA**GACTCTCCCAAGAGGCTCC
 CGTTCCCTCTCCAACCTCTTTGTCTCTTCTGCCCCAGTTTTCTTTATGGAGTACTTCTTTTC
 TCCGCTTTCTCTGTTTTCTCTTCTCTGCTCTCCCTCCCTCCACCTTTTTCTTCTCTTCC
 CAATTCCTTGCACTCTAACCAGTCTTGGATGCATCTTCTTCTTCCCTTCTCTTGTGTG
 TTCTTCTCTGTGTGTTTTGTGTGCCACATCTGTTTTTACCCCTGAGCTGTTTCTCTTTTT
 CTTTTCTTTCTTTTTTTTTTTTTTTTTTAAAGACGGATTCTCACTCTGTGGCCAGGCTGGAG
 TGCAGTGGTGCATCTCAGCTCACTGCAACCCCCGCTCTCTGGGTTCAGCGGATTCTCTCC
 CCCAGCCTCCCAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGACCCAGCCTGTTTCTCTTT
 TTCCACTCTTCTTTTTTCTCATCTCTTTTCTGGGTTCCTGTGCGCTTTCTTATCTGCCTGT
 TTTGAAGACCTTCTCTGTGTCTTGGGAGCCCTGAGACTTCTTCTCTCTTGCCTCCA
 CCCACTCCAAAGGTGCTGAGCTCACATCCACCCCTTGACGCCGTCCATGCCACAGCCCC
 CCAAGGGGCCCATTGCCAAAGCATGCCGCCACCCCTCGCTGTGCCTTAGTCAGTGTGTAC
 GTGTGTGTGTGTGTGTTTTGGGGGGTGGGGGGTGGGTAGCTGGGGGATTGGGCCCTCTTTCT
 CCCAGTGGAGGAAGGTGTGCACTGTACTTCCCTTTAAATTAAAAAACATATATATATAT
 ATTTGGAGGTCAATAATTTCCAATGGGCGGGAGGCATTAAGCACCGACCTGGGTCCCTAGG
 CCCGCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAGAATTTTTGCCAGGCTTACAGAACAC
 CCACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGCTGGATGCCTTTTATCCCAACTATTCT
 CTGTGGTATGAAAAA

FIGURE 20

```

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727
<subunit 1 of 1, 322 aa, 1 stop
<MW: 35274, pI: 8.57, NX(S/T): 1
MPVTIVTRTTITTTTSSSGLGSPMIVGSPRALTQPLGLLRLLQLVSTCVAFSLVASVGAWTG
SMGNWSMFTWCFCFSVTLLIILIVELCGLQARFPLSWRNFPITFACYAALFCLSASIIYPTTY
VQFLSHGRSRDHAIATFFSCIACVAYATEVAWTRARPGEITGYMATVPGLLKVLETFFVACI
IFAFISDPNLYQHQPALWCVAVYAICFILAAIAILLNLGECTNVLPPIPFPSFLSGLALLSV
LLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRRLAVAILTAINLLAYVAD
LVHSAHLVFVKV

```





































































Important features:

Transmembrane domains:

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192,
205-226, 235-255 and 294-312

N-glycosylation site.

amino acids 66-69

Glycosaminoglycan attachment site.

amino acids 18-21

FIGURE 21

FIGURE 22

MFLATLSFLLPFAHPFGTVSCEYMLGSPLSSLAQVNLSPFSPKPVHMDPNYCHPSTSLHLCS
 LAWSFTRLLHPPLSPGISQVVKDHTKPTAMAQGRVAHLIEWKGWSKPSDSPAALESASFSSY
 SDLSEGEQEARFAAGVAEQFAIAEAKLRAWSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPLG
 PHLQDLFTGHRFSRPVRQGSVEPESDCSQTVSPDTLCSSLCSELDGLLGSPARLASQLLGDE
 LLLAKLPPSRESAFRSLGLEAQDSLNSPLTESCLSPAEEEPAPCKDCQPLCPPLTGSWER
 QRQASDLASSGVVSLDEDEAEPEEQ

Signal peptide:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232,
 285-289, 324-328

Tyrosine kinase phosphorylation site.

amino acids 44-52

N-myristoylation site.

amino acids 17-23, 26-32, 173-179

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 11-22

FIGURE 23

GGTTCCTGGGCGCTCTGTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTCCTT
 GGCACCCCTCCTGCTCAGTGCACATTGTACACTTAACCCATCTGTTTTCTCTAATGCACGA
 CAGATTCCCTTTCAGACAGGACAACCTGTGATATTTAGTTCCCTGATTGTAAATACCTCCTAAG
 CCTGAAGCTTCTGTACTAGCCATTGTGAGCTTCAGTTTCTTCATCTGCAAAATGGGCATAA
 TACAATCTATTCTTGGCCACATCAAGGGATTGTTATTCTTTAAAAAAAACCAATACCAAG
 AAGCCTACA**ATG**TGTTGGCCTTAGCCAAAATTCTGTTGATTTCACGTTGTTTTATTCACTTCT
 ATCGGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAGTTT
 TTTAAACAATGGAAAATAAACCTATTCTTGGAAAGTGAAGCAAACCTTAACTCAGATAAA
 GAAAATATAACCACCTCAAATCTCAAGGCGAGTCATTCCCCTCCTTTGAATCTACCCAACAA
 CAGCCACGGAATAACAGATTTCTCCAGTAACATCATCAGCAGAGCATTCTTTGGGCAGTCTAA
 AACCCACATCTACCATTTCCACAAGCCCTCCCTTGATCCATAGCTTTGTTTTCTAAAGTGCCT
 TGGAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACC
 TGCTCTGTCTTCAGAAAACCTCACTTGGTCTTTGGTCAATGACACCGTGAAAACCTCTGATA
 TACAGTTCCATTACAGTTAGCATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCCCTG
 ATAGTGAACCAAGTGGATGGCTTACCACAAAACAGTGATAGCTTCACTGGGTTTACCCCTTA
 TCAAGAAAAACAACCTCTACAGCCTACCTTAAAAATTCACCAATAATTCAAACCTCTTTCCAA
 ATACGTCAGATCCCCAAAAAGAAAATAGAAAATACAGGAATAGTATTTCGGGGCCATTTTAGT
 GCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTGTGTGGAAAAAGGAAAAC
 GGATTCATTTTCCCATCGGCGACTTTATGACGACAGAAATGAACAGTTCTGCGATTAGACA
 ATGCACCGGAACCTTATGATGTGAGTTTGGGAATTCTAGCTACTACAATCCAACCTTTGAAT
 GATTGAGCCATGCCAGAAAGTGAAGAAAATGCACGTGATGGCATTCCATGGATGACATACC
 TCCACTTCGTACTTCTGTAT**TAGA**ACTAACAGCAAAAAGGCCGTTAAACAGCAAGTGTCATCTA
 CATCCTAGCCCTTTTGACAAATTCATCTTTCAAAGGTTACACAAAATTACTGTACAGTGGAT
 TTTGTCAAGGAGAAATCATAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCAT
 CCAAAGGTTTTCTTCTTACAATTTTGGCCATCCTGAGGCATTTACTAAGTAGCCTTAATT
 TGTATTTTAGTAGTATTTTCTTAGTAGAAAATATTGTGGAATCAGATAAACTTAAAGATT
 TCACCATTCAGCCCTGCCTCATAACTAAATAATAAAAATTTATTCACCAAAAAATTTCTAAA
 ACAATGAAGATGACTCTTTACTGCTCTGCCTGAAGCCCTAGTACCATAATTCAAGATTGCAT
 TTTCTTAAATGAAAATTGAAAGGGTGCTTTTAAAGAAAATTGACTTAAAGCTTAAAGAG
 GACATAGCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTCT
 TAGTACGTTATAATTTCTAGATCAGCACACACATGATCAGCCCACTGAGTTATGAAGCTGA
 CAATGACTGCATTACAGGGGGCCATGGCAGGAAGCTGACCCACCAGGAAAGTAATAGCT
 TCTTTAAAGTCTTCAAAGGTTTTGGGAATTTAACTTGTCTTAATATATCTTAGGCTTCAA
 TTATTTGGGTGCCTTAAAACTCAATGAGAATCATGGT

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732

><subunit 1 of 1, 334 aa, 1 stop

><MW: 36294, pI: 4.98, NX(S/T): 13

MLALAKILLISTLFYSLLSGSHGKENQDINTTQNIAEVFKTMENKPISLESEANLNSDKENI
TTSNLKASHSPPLNLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSFVSKVPWNA
PIADEDLLPISAHPNATPALSSENFTWSLVNDTVKTPDNSSITVSISSSEPTSPSVTPLIVE
PSGWLTTNSDSFTGFTPYQEKTTLQPTLKFNNKLFNPNTSDPQKENRNTGIVFGAILGAIL
GVSLLTLVGYLLCGKRKTDSFSHRRLYDDRNEPVLRLDNAPEFYDVSFGNSSYYNPTLND
SA
MPSEENARDGIPMDIPLRTSV

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
216

FIGURE 25

AACAGGATCTCCTCTTGCAGTCTGCAGCCCAGGACGCTGATTCCAGCAGCGCCTTACCGCGC
 AGCCCCAAGATTCACT**ATG**GTGAAAATCGCCTTCAATACCCCTACCGCGTGCAAAAGGAGG
 AGGCGCGGCAAGACGTGGAGGCCCTCCTGAGCCGCACGGTCAGAACTCAGATACTGACCGGC
 AAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGCTCCTCTGGGAGATGTATGCTTACTCT
 CTTAGGCCTTTCAATCATCTTGGCAGGACTTATTGTTGGTGGAGCTGCATTTACAAGTACT
 TCATGCCCCAAGAGCACCATTTACCGTGGAGAGATGTGCTTTTTTGATTCTGAGGATCCTGCA
 AATTCCCTTCGTGGAGGAGAGCCTAACTTCTCGCTGTGACTGAGGAGGCTGACATTGCTGA
 GGATGACAACATTGCAATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCCTGĈAG
 CAATTATTTCATGACTTTGAAAAGGGAATGACTGCTTACCTGGACTTGTTGCTGGGGAACGCG
 TATCTGATGCCCCCAATACTTCTATTGTTATGCCTCCAAAAAATCTGGTAGAGCTCTTTGG
 CAAACTGGCGAGTGGCAGATATCTGCCTCAAACCTTATGTGGTTCGAGAAGACCTAGTTGCTG
 TGGAGGAAATTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGA
 AAGTCCTTCGCGCTTCGTGCGAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAA
 ATGCTGGAAGATTAGACACTTCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAG**T**
AAGAGGCAACAGATAGAGTGTCCCTTGGTAATAAGAAGTCAGAGATTACAATATGACTTTAA
 CATTAAGGTTTATGGGATACTCAAGATATTTACTCATGCATTTACTCTATTGCTTATGCTTT
 AAAAAAAGGAAAAAAAAAAAAAACTACTAACCCTGCAAGCTCTTGTCAAATTTTAGTTAAT
 TGGCATTGCTTGTTTTTTGAAACTGAAATTACATGAGTTTCATTTTTTCTTTGCAATTTATAG
 GGTTTAGATTTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATCCATCC
 GTTGTTTTTTTTGTGTTGTTGTTTTTCTTTTCCTTTAAGTAAGCTCTTTATTCATCTTATG
 GTGGAGCAATTTTAAATTTGAAATATTTTAAATGTTTTTGAACTTTTTGTTAAATATA
 TCAGATCTCAACATTGTTGGTTCTTTTGTGTTTTTCATTTGTACAACCTTCTTGAATTTAGA
 AATTACATCTTTGAGTTCGTGTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAATT
 TTCATGAGACAGTCATTTTAACTAATGCAGTGATTCTTTCTCACTACTATCTGTATTGTGG
 AATGCACAAAATTTGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCACAA
 CCCTATAATAAATTTTACTCTATACAAAAA

FIGURE 26


</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828


<subunit 1 of 1, 263 aa, 1 stop


<MW: 29741, pI: 5.74, NX(S/T): 1


MVKIAFNTPTAVQKEEARQDVEALLSRTVRTQILTGKELRVATQKEGSSGRCLMTLLGLSF
 ILAGLIVGGACIYKYFMPKSTIYRGEMCFDSEDPANSLRGGEPNFLPVTEADIREDDNIA
 IIDVPVPSFSDSDPAIIHDFEKGMTAYLDLLLGNCYLMPLNTSIVMPPKNLVELFGKILASG
 RYLPQTYVVREDLVAVEEIRDVSNLGIIFYQLCNRKSFRLRRDRLLGFNKRAIDKCWKIR
 HFPNEFIVETKICQE


 **Type II transmembrane domain:**


 amino acids 53-75


 **N-glycosylation site.**

 amino acids 166-170

 **Casein kinase II phosphorylation site.**

 amino acids 35-39, 132-136, 134-138

 **N-myristoylation site.**

 amino acids 66-72, 103-109

Prokaryotic membrane lipoprotein lipid attachment site.


 amino acids 63-74

FIGURE 27

GGAGGAGGGAGGGCGGGCAGGCGCCAGAGCCAGAGCAGCCCGGGCACCAGCACGGACTCTCT
 CTTCAGCCAGGTGCCCCCACTCTCGCTCCATTGCGCGGGAGCACCCAGTCTGTACGCC
 AAGAACTGTGCTCTGGGGGCAC**CATG**GTTTCGGCGGCAGCCCCAGCCTCCTCATCCTTCTG
 TTGCTGCTGCTGGGCTCTGTGCTGTACCGACGCCGCTCTGTGCCCTGAAGGCCACGTT
 CCTGGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGCTCGTCGGCCTCCTCCCGAGCCTCC
 CGCCACCTTGACCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCACAACCTGGGG
 GGCCCATCACCCCCACCAACTTCTGGATGGGATAGTGGACTTCTTCCGCCAGTACGTGAT
 GCTGATTGCTGTGGTGGGCTCCCTGGCCTTCTGCTGATGTTTCATGCTGTGCGCGGTCA
 TCACCCGGCAGAAGCAGAAGGCCTCGGCCTATTACCCATCGTCTTCCCCAAGAAGAAGTAC
 GTGGACCAGAGTGACCGGGCGGGGGCCCCCGGGCCTTCAGTGAAGTCCCCGACAGAGCCCC
 CGACAGCAGGCCCGAGGAAGCCTGGATTCTCCCGCAGCTCCAGGCCGACATCTTGGCCG
 CCCACCCAGAACCTCAAGTCCCCACCAGGGCTGCACTGGGCGGTGGGGACGGAGCCAGGATG
 GTGGAGGGCAGGGGCGCAGAGGAAGAGGAGAAGGGCAGCCAGGAGGGGACCAGGAAGTCCA
 GGGACATGGGGTCCCACTGGAGACACCAAGGCGCAGGAGGAGCCGTCTCTTGTAGCCCCAG
 AGGGGGCTGTGGTGGCCGGTGAGGGCCAAAGGGAGCTGGAAGGGTCTCTCTTGTAGCCCCAG
 GAAGCCAGGGACCACTGGGTCCCCCGAAAGCCCTGTGCTTGACGAGTGTCCACCCAG
 GTGT**TAA**CAGTCTCCCGGGTGCCAGCCCTGACTGTGCGGCCCCCAAGTGGTCACTCCCC
 GTGTATGAAAAGGCCTTCAGCCCTGACTGCTTCTTGACACTCCCTCCTTGGCCTCCCTGTGG
 TGCCAATCCAGCATGTGCTGATTCTACAGCAGGCAGAAATGCTGGTCCCGGTGCCCCGGA
 GGAATCTTACCAAGTGCCATCATCCTTACCTCAGCAGCCCCAAAGGGCTACATCTTACAGC
 ACAGCTCCCCTGACAAAGTGAGGGAGGGCACGTGTCCCTGTGACAGCCAGGATAAAACATCC
 CCCAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCCAAACTACTTTTAAACA
 GCTACAGGGTAAATCTGCAGCACCCACTCTGAAAAATACTGCTCTTAATTTTCTGAAGG
 TGGCCCCCTGTTTCTAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTTAAATCCTC
 TCAAGCGCTCTCAAGCACCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGTGG
 GATCAGGTTGAATGAATGGAACCTTCTCTGTCTGGCCTCCAAGCAGCCTAGAAGCTGAGGG
 GCTGTGTTTGAGGGGACCTCCACCTGGGGAAGTCCGAGGGGCTGGGGAAGGGTTTCTGACG
 CCCAGCCTGGAGCAGGGGGGCCCTGGCCACCCCTGTGCTACACATTGTCTGGCAGCCTG
 TGTCCACAATATTCGTAGTCTCTCGACAGGGAGCCTGGGCTCCGTCCTGCTTTAGGGAGGCT
 CTGGCAGGAGGTCTCTCCCCATCCCTCCATCTGGGGCTCCCCAACCTCTGCACAGCTCT
 CCAGGTGCTGAGATATAATGCACAGCACAATAAACCTTTATTCCGGCTGAAAAAAAAAAAA
 AAAGA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852

><subunit 1 of 1, 283 aa, 1 stop

><MW: 29191, pI: 4.52, NX(S/T): 0

MVSAAAPSLILLILLLLGSPATDARSVPLKATFLEDVAGSGEAGSSASSPSLPPPWTFAL
SPTSMGPQPTTLGGPSPPTNFDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKA
SAYYPSSFPPKKKYVDQSDRAGGPRAFSEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSP
TRAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVPVETPEAQEEPCSGVLEGAVVAGE
GQGELEGSLLLAQEAQGPVGPPESPCACSSVHPSV

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 94-118

N-myristoylation site.

amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199,
211-217, 238-244, 242-248

FIGURE 29

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212

><subunit 1 of 1, 440 aa, 1 stop

><MW: 42208, pI: 6.36, NX(S/T): 1

MKFQGPLACLLALLCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
AGSKVSEALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
ADAVRGSWQGVPGHSGAWETSGGHGIFGSGGGLGGQGGQGNPGGLGTPWVHGYPGNSAGSFGM
NPQGAIPWQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
GSQSGSSGSGSNGDNNNGSSSSGSSSSGSSSSGSSSSGSSSSGSSSGNSGGSRGDSGSESSW
GSSTGSSSGNHGGSGGGNGHKGPCCEKPGNEARGSGESGIQGFRRQGVSSNMREISKEGNRL
GGSGDNYRGQSSSWGSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDO
RSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 31


GACCCGTCCTCCGTCCTGGATTGTGCGGACTCTGCTGCAGCGAGGGCTGCAGGCCCGCGGGCGGTGCTCACCG
 TGCCCTGGCTGGTGAGATTCTCTCTCTTGGCTGACCATTGTGTTCCTTGTCTGGAATATTACCGGGACATCTTCA
 CTCTCCCTGTGCGCTGCACCGGAGCTTGGTGTGTGCGCAGGAGAGTGAAGGGAAGATGTGTTTCTCTGGAACAAGC
 TGCTGCTACTTGTCTCTCGGGCTGGCTTTTCCAGATTCCACAGCTCCTGAGGACTTGTCTTTCTTGGAAAGAGG
 GTCCCTCATATGCTTTGAGGTGGACACAGTAGCCCCAGAGCATGGCTTGGACAAATGCGCTGTGGTGGACCAGC
 AGCTGCTCTACACCTGTGCGCCCTACATCGGAGAGCTCCGGAAATCTGCTCGCTTGTGGGTGTTCAGGCAGTAGTG
 GACGGAGTGGGGGCTTCATGAGGAAAATCACCCGCCACTACCACAGCCCTGGGAGCCCGAGCTTCCAGAGCA
 GCCAGGGGTGCGAGGCACAGCTCGCCAGGGCCTTTTCCACAACAGCGCCCTCCTTGCGCCGAGCCGCTAGAGT
 TCGTGGCAGAAAGAAATTTGGATCAAACTGTGTCAAAACATATCAAGGCTACACTGGTGGCAGATCTGTCGCCAGG
 CAGAGTCACTTCTCAAGAGCAGCTGGTGACACAGGGAGAGGAAGGGGAGAGCCAGCCAGCTGTTGGAGATCT
 TGTGTTCCAGCTGTGCCCTCACGGGCCAGGCATGTGCCCTGGGCGGGAGTTCTGTCAAGGAAGAGCCCTG
 GGGCTGTGCGGGCGCTGCTTCCAGAGGAGACCCCGGAGCCGTTCTGAGCAGTGCAGAGAACATTGCTGTGGGGC
 TTGCAACAGAGAAAGCCTGTGCTTGGCTGTCAAGCAACATCACAGCACTGATCAGGAGGGAGGTGAAGCAGCAG
 TGAGTGCACACTTCGAGCCAGGGTCTGAACCTGCTGCCCGGGGAGCGGAGGGGCTGCTCCCGCGCTTCAAGC
 GTGCTCTCTTGGCCGTGGGGCCAGGGACCTGACGAGGGAGTCTCCCGAGCATCTGGAACAGCTCTTAGCC
 CAGCTGGGCCAGACGCTCGGTCGCCCGAGTTCTGTGCCCACTGCTGAGCAGCATCTGGCAAGTCTCTGTGTG
 GAGTTAGCTTCCCTCCTCTGTGACAGTCAAAATTCCTATCTTAGGCCCCCGGACAGTACAGGCTGGAGAGAGG
 CAGGCTCGAAGGCTTCTGCACATGCTGCTTCTTCTTGTGGAAGGAGACTTTCAGGGCCGGTTCGCGTGAGCTG
 CTGCTGAGCCCAAGAAATGTGGGGCTTCTGCGACAGCACAGGCCAAGGAGTGGGAGTCTGCTGCTATTCTTGCTA
 CGGGAGCTGCTGGACAGGCTCTGATGGGACGGATGGAGTAGAGGCTTGCCTGGGACGCTTCCACAGGCCCAG
 TGCCAGGGGACTTGTCTGAAGAAATAGCACACTGTCTAATCTGTTCTAGCCGAGCCCACTTGCAGGAACCC
 CAGCTAAGACCTGTGAGTTGGTGACGCCAAACCGGGGCATGTGCTGGCCAGAGCTAGGGCTGAGAAGTGGCC
 CTGCGCTTGGGACTGCACCGAAGCCTGGACCCCGCCTCACGAGGAGGCCAAGTGCCCAATGCAGACCCCTCA
 TGGTTGGGCTGTAGCTGGGTCTACAGTCAGACTTCTGCTCTAAGGTTGTCACTGTGGCATCCACACCGCA
 ATCTAGAGGAAGGAGAGTGGCGCTGATTTGGGATTATGGCAAAAAGTCCAGAGATGCCAGTCTCTGGAGTAGAA
 TGGTGGTGTGTTGTTTATCTCTTGGATACTAAATGAAATGAGGTGTGGGGCTTGTGGGCTTGTCAACACAGAAATCAAGCCT
 CATTTGCTATCCAGCATCTCTTAAAACTTTGAGTCTTGAAGTTCATGACAGGGCAAAATGACTCTGCTTAAC
 TTATGAAGAAAGTTAAACATGAATCTTGGGAGTCTACATTTCTTATCACCAGGAGCTGGACTGCCATCTCCTT
 ATAAATGCTTAACACAGCGCGGTCTGTTGGCTCATGCTGTAATCCAGCACTTTGAGAGGCTCAGGTCGCGG
 GACTGCTGAGGTGAGGAATTCAGACACAGCTGGCCAACTGGCAAAACCCCATCTCTACTAAAAATAAAAAA
 TTATTAGCTGGGCTATGTTGTTGTGTGCTGTAATCCAGCTACTCAGGAGGATGAGGCAGGAGACCTGCTTGAAC
 CTGAGGTGGAGGTTGCAGTGAGCCGAGGTGCGCAACCTGCACTCCAAGTCTGGGTAACAGAGCGAGACTTTCTAG
 AAAAGCCTAACAAACAGATAAGGTAGGACTCAACCACTGAAACCTGACTTTCCCCCTGTACTCTCAGGCCCTG
 TGCAGGTAGTAACCTTTGAGACCTCTCCCTGACCAGGGACCAAGCACAGGGCATTTAGAGCTTTTAGAATAAA
 CTGGTTTTCTTTAAAAAAGGCGCGCCGCTTT
 TTTTAAAAAGGGCTTTTATTAATAATCTCCCAACAGTGGCTGCTGCAATCTGCCACAGCTC
 TGGGGCGTCTCTGTAGGGAAGGGCTGTTTTCTCCTGAGGCGGGGCTGGGCTTGTCCATGGGTCCGCGAGCTG
 CCGCTGCTTGGCGCCCTGGCGTGTCTAGCTGCTTCTTGGCGGCACAGAGCTGCGGGGCTTGGGGGCCACGG
 AGCTAACAGAGCAGCTCTGGTGACGGGTGGAGGCTGTCTCTTAAACGACACCTCAGGTGCTCTGAGATGCTG
 GGTACCCCTGAGGCTGAGGAGCAGCTGTGGCCGGTGTCTCTCYTAGCGCTCTCTGGGGAATCAAGCTC
 GGGCCCTCTTTGCAAGAGCCAGGATGGGTGGGTGTGGGGGACTCATGGGAAATGGCTGAGGAGCTACGTGT
 GAAGAGGCGCGGGTTGTTGGCTCAGCGGCCCTGAGCGCCTCTCTCTGAGCCTCAGTTTCCCTTTCCGCTA
 ATGAAGAACATGCGCTCTCGTGTCTCAGGCTATTAGGACTTGCCCTCAGGAAGTGGCTTGGACAGGAGCTCAT
 GTTATTCTTCAACATGTCTCGCAGCTGTGGCTGGGACGTCATGGAATGGCCATGTCCCTCTGCTGGTGGAC
 GTGCGGTCGGGAGTGGCGACCCAGAGCGGGGCGAGACGTGGCTGGGGTGGGGGAGGGGCGCCCGGAGG
 CCTCACAGGAATGGGCTCCGCCACACAGGACAGGGCGGGCTCCCGCCGCGCGCCGACACCTGTCAGG
 GGCGGTAGCAAAAGTGAAGTGTGGCTTGGGCTGCTGCGCAGCAGGTAGCCCTGATGCACTGCGGCAGCGCG
 TCGTCCGCGAGCTGGAAGCAGCGCCGCTCCACAGCACGAACGCGGTGGCCCT


FIGURE 32


MCFLNKLLLLLAVLGWLFQIPTVPEDLFFLEEGPSYAFEVDTVAP EHG LDNAPVVDQQLLYTC
 CPYIGELRKLLASWVSGSSGRSGGFM RKITPTTTTSLGAQPSQTSQGLQAQLAQAFFHNQPP
 SLRRTVEFVAERIGSNVCVKHIKATLVADLVRQAESLLQEQLVTQGEEGGDP AQLLEILCSQL
 CPHGAQALALGREFCQRKSPGAVRALLPEETPAAVLSSAENIAVGLATEKACAWLSANITAL
 IRREVKA AVSRTLRAQGPEPAARGERRGCSRA


Signal peptide:


amino acids 1-18


 **N-glycosylation site.**


 amino acids 244-248

 **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

 amino acids 89-93

 **Casein kinase II phosphorylation site.**

 amino acids 21-25, 167-171, 223-227

 **N-myristoylation site.**

amino acids 100-106, 172-178, 207-213

Microbodies C-terminal targeting signal.

amino acids 278-282

FIGURE 33

TCCCTTGACAGGCTCTGGTGGCTGGTTCTGGGGTCTACTGAAGGCTGTCTTGATCAGGAACTG
 AAGACTCTCTGCTTTTGGCCACAGCAGTTCTTGCAGCTTCTTTGAGGTGTGAACCCACATCCC
 TGCCCCCAGGGCCACCTGCAGGACGCGGACACCTACCCCTCAGCAGACGCGCGGAGAGAA**ATG**
 AGTAGCAACAAAGAGCAGCGGTGACGAGTGTCTGTGATCCTCTTTGCCCTCATCACCATCCT
 CATCCTCTACAGCTCCAACAGTGCCAAATGAGGCTTCCATTACGGCTCCCTGCGGGGCCGTA
 GCCGCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGC
 AACAAAGACTGCCCTCTCGGTGCCACCAGTGTGTATTGTGACGAGCTCCAGCCACCTGCCT
 GGGCACCAGCTGGGCCCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGCAC
 CCACCACTGGCTACTCAGCTGATGTGGGCAACAAGACCACCTACCCGCTCGTGGCCCATTC
 AGTGTGTTCCGCGTGCTGAGGAGGCCCCAGGAGTTGTCAACCGGACCCCTGAACCCGTGTT
 CATCTTCTGGGGGCCCCGAGCAAGATGCAAGAAGCCCCAGGCGAGCTCGTGCGTGTATCC
 AGCGAGCGGGCCTGGTGTTCCTCAACATGGAAGCATATGCCGTCTCTCCCGGCCCATGCGG
 CAATTTGACGACCTCTTCCGGGGTGAGACGGGCAAGGACAGGGAGAACTCTCATTCGTGGTT
 GAGCACAGGCTGGTTTACCATGGTGATCGCGGTGGAGTTGTGTGACCACTGCATGTCATAG
 GCATGGTCCCCCAACTACTGCAGCCAGCGGCCCGCCTCCAGCGCATGCCCTACCACTAC
 TACGAGCCCAAGGGGCCGAGCAATGTGTCACTACATCCAGAATGAGCACAGTCGCAAGGG
 CAACCCACACCGCTTCATCACCGAGAAAAAGGGTCTTCTCATCGTGGGCCAGCTGTATGGCA
 TCACCTTCTCCACCCCTCCTGGACCT**TAGG**CCACCCAGCCTGTGGGACCTCAGGAGGGTCAG
 AGGGAAGCAGCTCCGCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTG
 CTGGAGTGTCTCCAGCCAATCAGGGCCTTGAGGAGGATGTATCTCCAGCCCATCAGGGCC
 TGGGGAATCTGTTGGCGAATCAGGGATTTGGGAGTCTATGTGGTTAATCAGGGGTGTCTTTCT
 TTGTGCAGTCAGGGTCTGCGCACAGTCAATCAGGGTAGAGGGGGTATTTCTGAGTCAATCTG
 AGGCTAAGGACATGTCTTTCCCATGAGGCCTTGGTTCAGAGCCCCAGGAATGGACCCCCCA
 ATCACTCCCACTCTGCTGGGATAATGGGGTCTGTCCCAAGGAGCTGGGAACCTTGGTGTG
 CCCCCTCAATTTCCAGCACAGAAAGAGAGATTGTGTGGGGGTAGAAGCTGTCTGGAGGCC
 GGCCAGAGAATTTGTGGGGTTGTGAGGTTGTGGGGCGGTTGGGAGGTCCAGAGGTGGGA
 GGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTGGACAAACCCCTTCCCCCTCTCTGGG
 CACCCCTCTGCCACACAGTTCAGTGCGGAGTCTGAGACCCCTTCCACCTCCCCTACAA
 GTGCCCTCGGGTCTGTCTCCCCGTCTGGACCCCTCCAGCCACTATCCCTTGTCTGGAAGGCT
 CAGCTCTTTGGGGGGTCTGGGGTGACCTCCCCACCTCCTGGAAAACTTTAGGGTATTTTTCG
 GCAAACCTCCTTCAGGGTTGGGGGACTCTGAAGGAAACGGGACAAACCTTAAGCTGTTTTCT
 TAGCCCTCAGCCAGCTGCCATTAGCTTTGGCTCTTAAAGGGCCAGGCCTCCTTTTCTGCCCT
 CTAGCAGGGAGGTTTTCACACTGTGTGGAGGCGCCTTTGGGGCTGCCCTTTGTCTGGAGTCA
 CTGGGGGCTTCCGAGGGTCTCCCTCGACCTCTGTGCTCTGGGATGGCTGTCTGGGAGCTGT
 ATCACTTGGTCTGTCTCCCTGGCTCTGTATCAGGCACCTTATTAAAGCTGGGCTCAGTGG
 GGTGTGTTTGTCTCTGTCTTCTGTGAGCCTGGAAGGAAGGGCTTCAGGAGGAGGCTGTGA
 GGCTGGAGGGACCAAGTAGGAGGAGCCAGCAGCTAGCCATTGCACACTGGGGTGATGGGTGG
 GGGCGGTGACTGCCCCAGACTTGGTTTGTGAATGATTGTACAGGAATAAACACACCTACGC
 TCCGGA

FIGURE 34

MSSNKEQRSADFVILFALITILILYSSNSANEVFHYGSLRGRSRRPVNLKKWSITDGYVPIL
 GNKTLPSRCHQCIVSSSSHLLGTLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAH
 SSVFRLRRPQEFVNRTPETVFIFWGPPSKMQKPQGSILVRVIQRAGLVFPNMEAYAVSPGRM
 RQFDDLFRGETGKDREKSHSWLSTGWFTMVIAVELCDHVHVYGMVPPNYCSQRRLQRMPIYH
 YYEPKGPDECVTYIQNEHSRKGNNHHRFITEKRVFSSWAQLYGITFSHPST

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 9-31 (type II)

N-glycosylation site.

amino acids 64-68, 115-119

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 50-54

Casein kinase II phosphorylation site.

amino acids 3-7, 29-33, 53-57, 197-201

Tyrosine kinase phosphorylation site.

amino acids 253-262

N-myristoylation site.

amino acids 37-43, 114-120, 290-294

FIGURE 35

GTTTCATAGTGTGGCGTCTTCTTAAAGGAAAAACACTAAATGAGGAACCTCAGCGGACCGGGAGCGACGAGCTT
 GAGGGAGAGCATCCCTAGCTTTGGGCGCAGAGGGGGCAGGCTGAAGCCGAGTGGCCCGAGGTGTCTGAGGGGCTG
 GGCAAGGTCGAAGNGTTTCAGAACAGCTTCTCGAACCCATGACCCATGAATCTTGTGCACATTTATACGGT
 CTGAGGCTACAGCTCGAACTACAGAAAGTGGAGTCTGGCCAGGAGCGCAGTATCTCTTTGTGTGAOCTTGGC
 GGCTATGGAGCACTTGGCTTCAGACCTTTGTGATACACCA**TAGCT**GGCTGGGACAGTACGGCGTGGAGAGGAATG
 AGGCGTGGGCTCACATGGCTTCCCTCTCTAGCCACAGCAGGCTGCTTTGCTGACTTGAACAGAGTCCCTCAG
 GTACACCTCCAGCCTGGCTCCACCGTCCAGAGCCCGGAGGCACTGTGATCTTGGGCTGGCTGGTGAACCTCCA
 AGGATGAATGAACCTGGGCGCTGAATGGAAGAGAGTGAATGGCTCGGATGAGCTCTGGGTGTGCTCTCATCAC
 CACGGGACCTCGTCACTCATGCCCTTAAACAACCACTGTGGGACGGTACCAAGTGTGTGGCCGGATGCTCTGGC
 GGGGCTGTGGCCAGCGTGGCAGGCACTGTGACACTAGCCAACTCTCCAGGACTTCAAAGTATAGTGTGAGCAGGTG
 ATTGAAGTGGATGAGGGAAACACAGCAGTCACTTGGCTGCCACCTGGCTGAGAGCCACCCCAAAGCCAGGTCCGG
 TACAGGCTCAACACAGAGTGGCTGGAGGCTCCAGAGGTAACATCTGATCATGCCCTCAGGGAACCTTCAGATTT
 GTGAATGCCAGCCAGGAGGACAGGGCATGTACAAGTGTGACAGCTACAAACCCAGTGACCCAGGAAGTGAAGAAC
 TCCGGCTCCAGCGACAGGCTACGTGTGCGCCGCTCCACCGCTGAGGCTGCCCGCATCATCTACCCCCCAGAGGCC
 CAAACCATCATCGTCAACAAAGGCCAGAGTCTCATTTGGAGTGTGTGGCCAGTGGAAATCCCAACCCCCACGGGTG
 ACCTGGGCCAAGGATGGGTCCAGTGTCCACGGCTACAACAAGACCGGCTTCTCTGTGAGCAACCTCTCATCGAC
 ACCACCGAGGAGGAGGACTCAGGCACTACCGCTGCATGGCCGACAAATGGGTGGGCGAGCCGGGGCAGCGGTG
 ATCTCTACAAGTGTCCAGSTGTTTGAACCCCTGAGGTACACCATGGAGCTATCCCAAGTCTGCTCCCTGGGGG
 CAGAGTGCCAAAGCTTACCTGTGAGTGTGGTGGGAACCCCGCCCTCCGCTGTGGCTGGGCTGAGGATGCTGTGCC
 CTCTATCTCCAGACCGCGCTCGCGCTCTCCAGGGCCCTGCGCTGCTCAGCATGGGCGCTGAGGACAGAGCG
 STCTACCAATGCACTGGCAGAGACAGGTTGGGAGCGCCCATGCCGTAGTCCAGCTGGGAGCTCTCAGGCCAAGC
 ATACCCCAAGAGCTATGGCAGGATGCTGAGCTGGCTACTGGCACACCTCTGTATCACTCCAAACTCCGCCAAGC
 CCTCGACAGATGCTGAGGGGGCAACCGGCGCTCCCGAGACCCCAACCTCAGTGGGGCTGCTTCCCCAGAGTGT
 CCAGGAGAGAGGGGGCAGGGGGCTCCCGCGAGGCTCCCATCATCTCTCAGCTCGGCCCGCACTCCAGAGCAGAC
 TCTATATGAACCTGGTGTGGCGGCTCGGCATGAGGCGAGTGGCCGGGCGCAACTCTCTACTATGTGTGAACAC
 CGCAAGCAGGTACAAATTTCTCTGACGATTGGACCATCTCTGGCATTCCAGCAACAGCAGCCGCTGACCCCTC
 ACAGACTTTCGCCCGGAGCTTGTATGAAGTGGAGATGGCAGCTTACAACCTGTGCGGGAGAGGGGCCAGACGCC
 ATGGTCACTTCCGAAGTGGAGCGGGGCCAAACCCGAGATCATGGCCAGCAAGAGCAGAGATCCAGAGAGAC
 GACCTTGGAGCCAGTCCCGAGCGAGCCAGCCAGACACCGGCGCTCTCCCCCAGAGACTCCCGAGAGG
 CCCACCTATCCACGGGCTCCGAGACCTCAGTGTACCTGACCTGGATTCGCCCTGGGAATGTGTGGTTTCCCAATC
 CAGTCCCTCCGTGTGGATGACAGAAAGCTAAAGAAAGTGGGAGACTGGATTCTGGGCAACCGCGCCATCCCCCA
 TCGCGGCTGTCCGTGGAGATCAGCGGCCCTAGAGAAAGGACACCTCTTACAAGTTTCGAGTCCGGGCTCTGAACATG
 CTGGGGAGAGGAGCCCGCAGCGCCCTCTCGGCCCTACGTGGTGGTCTGGGCTCAGCGGTCGCTGTGTACGAGATG
 CCGCTGGCAGGCTTATATACCTTACGGATGGGCTCATGAGACCATCATGCTCAAGTGAATGTACATC
 CCAGCAAGTAAACACACACCCCAATCCATGGCTTTTATATCTATATACGACCCAGACAGATGACATATGATGT
 GACTACAGAAAGATATGGTGGAAAGGGGACAAGTACTGGCACTCCATCAGCCACTCGACCGAGCACTCTCTAC
 GAGATTAAGATCGAGTGTCTCAATGAAGGAGGGGAGAGCAGTTCAGCAACGTGATGATCTGTGAGACCAAAGCT
 CGGAAGTCTTCTGGCCAGCTGGTGTGACTGCCAACCTCTGGCCCCACACAGCGCCGCTTCTCTGAAGAC
 ATAGAGCGGCGGGTGGGCATCGGGCCATGGTGGCTCGCTCCAGCGACCTGCCCTATCTGTGTCTGGGGCTGTCT
 CTGGGCTCCATGCTTCTCATCTCATCTGTCACCTTCACTCCCTTCTGCTTGTGGAGGGCTGATGTAAGCAAAACAT
 ACAACAGACCTGGGTTTTCTCGAAGTGGCTTCCACCTCTCTGCCGTATATCTATGGTGCCATTTGGGAGGACTC
 CCAGGCCACCGGCGAGTGGGAGCCCTACCTCAGTGGCATCAGTGGAGCGGCGCTGTGCTAATGGGATCCAGATG
 AATAGGGGCTGCCCTCGGCTCGAGTGGGCTACCCGGGCTGAAGCCCGAGCAGCATGCCCAGGCGAGCTTCAG
 CAGCAGAGTGAACAGCAGCGCTGTCTGAGGCAGACCCATCTTGGCAATGGATATGCCCCAAATGACCCCAAGT
 ACGAGGGGTCCCAAGTCTAGCCCGGACAGGGGCTCTTTCTTATACACATGCCCGAGGCTCCACTCAGCAGCTGT
 CTGAGCCCCATCACGACTGCTGCAAGCCGAGGAGCAGCTGCTGTCTGGGCGCTCAGGGGTGAGGAGAGCC
 CCGCAGCTCTGTCTCTGGAGATGTGGGCCCTTCCATTTCTCAGGCGCTCTGCTGTGGTCTGTGGGCTGTG
 CCAAGTGAAGAGTGGACATCTCTGACTCTGCGAAGTGAAGTGGAGCAGCTGGTGTCCCCAGGCTGAAGG
 GCTTAAGTAGAGAGCAACCTGGAATGAGCTCTCCCCGGGGCGACTGGTGGGTGTCTTTTGAACACCACTCT
 CTCACAT**TAGG**CAGAAGCTGATATCCAGAAAGACTATATATTGTTTTTTTTTAAAAAAAAGAGAAAAA
 AGAGACAGAGAAAAATGGTATTTATTTTCTTATATAGCCATTTATATATATGCACTTTGAATTAATGTA
 TATGTTTTTAAATTTCTGGAGAGACATAGGAGTCTTACCCGTGAGGTTGGAGAGGAAAAATAAGAGGTGCCA
 CCTAACAGGAGTCAACCCAGAAAGCACCGCAGGCTGGCGGGGACAGACTCTAACCTGGGGCTCTCGAGTG
 CGAGGCGAGGCTCGAGGAGGCCCCACAGATAAGCTGGCAAGAGGAAGGATCCAGGCACATGGTTCATCAGGACA
 TGGAGGAAACCAAGAGGGCCAGTATCACAGCTGGAGACACCCACACAGATGGCTGGGTGCTACGGGAA
 CAATTTCTCTAAGATGCCATGAGACAGACCAAGATGTGTACAGCACTATGACATCTTAAATAACCTTCAGAA
 CAATAATTCGGTGGCAACATATCTCTGTAAAAACAAACACTGTAACTTCAATTAATATGTTTAACTTCTCCCTGTAA

FIGURE 36

MLRGTMTAWRGMREVTIACLLLATAGCFADLNEVPQVTVQPASTVQKPGGTVILGCVVEPP
 RMNVTVRLNGKELNGSDDALGVLITHGTLVITALNNHTVGRYQCVARMPAGAVASVPATVTL
 ANLQDFKLDVQHVEVDEGNTAVIACHLPESHKPAQVRYSVKQEWLEASRGNYLIMPSGNLQ
 IVNASQEDEGMYKCAAYNPVTQEVKTSGSSDRLRVRRSTAEAAARIYPPEAQTIIVTKGQSL
 ILECVASGIPPRVTWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGA
 AVILYNVQVFEPPEVTMELSQLVPIWQSAKLTCCEVRGNPPPSVLWRNAVPLISSQRLRLS
 RRALRVLSMGPEDEGVYQCAENEVGSAAHAVVQLRTSRPSITPRLWQDAELATGTPPVSPSK
 LGNPEQMLRGQPALPRPPTSVGPASPKCPGEGKGQGAPEAPIILSSPRTSKTDSYELVWRPR
 HEGSGRAPILYYVVKHRKQVTNSSDDWTISGIPANQHRLLTLRLDPGSLYEVEMAAYNCAGE
 GQTAMVTFRTGRRPKEIMASKEQQIQRDDPGASPGSSSQPDHGRLSPPPEAPDRPTISTASE
 TSVYVTWIPRGNGGFPIQSFRVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGTSYKFRV
 RALNMLGESEPSAPSRPYVVSIGYSGRVYERPVAGPYITFTDAVNETTIMLKWMIIPASNNNT
 PIHGFIYYRPTSDNDSYKKDMVEGDKYWHSISHLQPETSYDIKMQCFNEGGESEFSNVM
 ICETKARKSSGQPGRLPPPTLAPPQPLPETIERPVGTMAMVARSSDLPYLIVGVVLGSIVL
 IIVTFIPFCLWRAWASKQKHTTDLGFPRLSALPPSCPYTMVPLGGLPGHQASGQPYLSGISGRA
 CANGIHMNRGCPSAAVGYPGMKPQQHCPGELQQQSDTSSLLRQTHLNGYDPQSHQITRGPK
 SSPDEGSFLYTLRDDSTHQLLQPHHDCCQRQEQPAAVGQSGVRRAPDSVPLEAVWDPFPHSG
 PPCCCLGLVPVEEVDSPDSCQVSGGDWCPQHVPVGAAYVQEPGMQLSPGPLVRVSFETPPLTI

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 16-30 (type II), 854-879

FIGURE 37

CGGGAGGCTGGGTCGTCATGATCCGGACCCCATTTGTCGGCCTCTGCCCATCGCCTGCTCCTC
 CCAGGCTCCCGCGGCCGACCCCGCGCAAC**ATG**CAGCCACGGGCCGCGAGGGTTCCCGCGC
 GCTCAGCCGGCGGTATCTCGGGCGTCTGCTGCTCCTGTACTGCTGCTGCTGCTGCGGCAGC
 CCGTAACCCGCGCGGAGACCAGCCGGGCGCCCCAGAGCCCTCTCCACGCTGGGCTCCCC
 AGCCTCTTACCACGCGGGTGTCCCAGCGCCCTCACTACCCAGGCCTCACTACGCCAGG
 CACCCCAAAACCTGGACCTTCGGGGTCGCGCGCAGGCCCTGATGCGGAGTTTCCCACTCG
 TGGACGCCACAATGACCTGCCCCAGGTCCTGAGACAGCGTTACAAGAATGTGCTTCAGGAT
 GTTAACCTGCGAAATTTAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGT
 GGGTGCCAGTTCTGGTCAGCCTCCGTCTCATGCCAGTCCCAGGACCAGACTGCCGTGCGCC
 TCGCCCTGGAGCAGATTGACCTCATTCACCGCATGTGTGCCTCTACTCTGAACTCGAGCTT
 GTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAGCTGGCCTGCCTCATTGGCGTGNAGGG
 TGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCT
 AACTGACACTTACCTTCACCTGCAGTACACCATGGGCAGAGAGTTCCACCAAGTTCAGACAC
 CACATGTACACCAACGTCAGCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAA
 CCGCCTGGGCATGATGATAGATTGTCTATGCATCGGACACCTTGATAAGAAGGGTCTTGG
 AAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGTGCCAGAGCTGTGTGTGACAATTTG
 TTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGACACT
 GTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTG
 ACCACATCAGGGCAGTCATTGGATCTGAGTTCATCGGGATTGGTGGAAATTATGACGGGACT
 GGCCGGTTCCTCAGGGGCTGGAGGATGTGTCCACATACCCAGTCTGTAGAGGAGTTGCT
 GAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTGTCCTTCGTGGAAACCTGCTGCGGGTCT
 TCAGACAAGTGGAAAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCGTGGAGGCTGAGTTT
 CCATATGGGCAACTGAGCACATCCTGCCACTCCACCTCGTGCCTCAGAATGGACACAGGC
 TACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCTGGAGGTCTCAAATGCCT
 CCCCATACCTTGTTCAGGCCTTGTGGCTGCTGCCACCATCCCAACCTTACCCAGTGGCTC
 TGCT**TGA**CACAGTCGGTCCCCGAGAGGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCT
 AGTTCAATTCACAAGCATATGCTGAGAATAAACATGTTACACATGGAAAA

FIGURE 38

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817

<subunit 1 of 1, 487 aa, 1 stop, 2 unknown

<MW: 53569.32, pI: 7.68, NX(S/T): 5

MQPTGREGSRALSRRYLRRLLLLLLLLLLLLLQQPVTRAETTPGAPRALSTLGSPSLFTTPGVPS
 ALTTPLGLTTPGTPKTLDLRGRAQALMRSFPLVDGHNDLPQVLRQRYKNVLQDVNLRNFSHGQ
 TSLDRLRDGLVGAQFWSASVSCQSQDQTAVRLALEQIDLIHRMCASYSELELV TSAEGLNSS
 QKLACLIQVXGGHSLDSSLSVLRSFYVLGVRYLTLTFTCSTPWAESSTKFRHHMYTNVSGLT
 SFGEKVVEELNRLGMMIDLSYASDTLIRRVLVSQAPVIFSHSAARAVCDNLLNVPDDILQL
 LKNGGIVMVTL SMGVLQCNLLANVSTVADHFDHIRAVIGSEFIGIGNYDGTGRFPQGLEDV
 STYPLVIEELL SRXWSEELQGVLRGNLLRVFRQVEKVREESRAQS PVEAEFPYQQLSTSCH
 SHLVPQNGHQATHLEVTKQPTNRVPWRSSNASPYLVPGLVAAATIPTFTQWLC

Important features of the protein:

Signal peptide:

amino acids 1-36

Transmembrane domain:

amino acids 313-331

N-glycosylation sites.

amino acids 119-122, 184-187, 243-246 and 333-336

N-myristoylation sites.

amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359, 357-362, 394-399, 427-432 and 472-477.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 136-146

FIGURE 39

TGCTAGGCTCTGTCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACCCACAGA
 TCCCTCTATGACTGCAATGTGAGGTGTCCGGCTTTGCTGGCCAGCAAGCCTGATAAGC**ATG**
 AAGCTCTTATCTTTGGTGGCTGTGGTGGGTGTTTGCTGGTGCCCCAGCTGAAGCCAACAA
 GAGTTCTGAAGATATCCGGTGCAATGCATCTGTCCACCTTATAGAAACATCAGTGGGCACA
 TTTACAACCAGAATGTATCCCAGAAGGACTGCAACTGCCTGCACGTGGTGGAGCCCATGCCA
 GTGCCTGGCCATGACGTGGAGGCCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAG
 CACCACCACCATCAAGGTCATCATTGTCTATCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCT
 ACATGGCCTTCCTGATGTGTGGTGGACCCTCTGATCCGAAAGCCGGATGCATACACTGAGCAA
 CTGCACAATGAGGAGGAGAATGAGGATGCTCGCTCTATGGCAGCAGCTGCTGCATCCCTCGG
 GGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGCCCAGCAGCGGTGGAAGCTGC
 AGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGC**TAG**ATGGGCTGG
 GTGTGGTGGGTCAAGGCCCCAACACCATGGCTGCCAGCTTCCAGGCTGGACAAAGCAGGGGG
 CTA¹CTTCTCCCTTCCCTCGGTTCCAGTCTTCCCTTTAAAGCCTGTGGCATTTTTCCTCCTT
 CTCCCCTAACTTTAGAAATGTTGTA²CTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCT
 GATCTCTGTTGCTTCTTGGGTCTTTGGGGTTGAAGGGAGGGGGAAGGCAGGCCAGAAGGGA
 ATG³TGGAGACATTGAGGCGGCCCTCAGGAGTGGATGCGATCTGTCTCTCCTGGCTGCACCTTTG
 CCGCCTTCCAGCTCTGAGTCTTGGGAATGTTGTTACCTTGGAAAGATAAAGCTGGGTCTTCA
 GGA⁴ACTCAGTGTCTGGGAGGAAAGCATGGCCAGCATTAGCATGTGTTCTCTTCTGCAGTG
 GTTCTTATCACCACCTCCCTCCCAGCCCCGGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTG
 AGGACAGCTCTGATGGGAGAGCTGGGCCCCCTGAGCCCACTGGGTCTTCAGGGTGCAC⁵TGGA
 AGCTGGTGTTCGCTGTCCCTGTGCACTTCTCGCACTGGGGCATGGAGTGGCCATGCATACT
 CTGCTGCCGGTCCCTCACCTGCACTTGAGGGGTCTGGGCAGTCCCTCCTCTCCCCAGTGTC
 CACAGTCACTGAGCCAGACGGTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGAA
 CACCACAGCCCTGTACTTGGGTGCGCTCTTGTCCTGAACTTCGTTGTACCACTGCATGGA
 GAGAAAATTTGTCTCTTGTCTTAGAGTTGTGTGTAATCAAGGAAGCCATCATTA⁶AATTG
 TTTTATTTCTCTCA

FIGURE 40




</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278

<subunit 1 of 1, 183 aa, 1 stop



<MW: 20574, pI: 6.60, NX(S/T): 3

MKLLSLVAVVGCLLVPPAEANKSSDIRCKCICPPYRNISGHIYNQNVSQKDCNCLHVVEPM
PVPBGHDVEAYCLLCECRYEERSTTTIKVIIIVYLSVVGALLYMAFLMLVDPLIRKPDAYTE
QLHNEEENEDARSMAAAAASLGGPRANTVLERVEGAQQRWKLQVQEQRKTVFDRHKMLS

Important features:**Signal peptide:**

 amino acids 1-20



Transmembrane domain:

 amino acids 90-112


N-glycosylation sites.




 amino acids 21-24, 38-41 and 47-50



FIGURE 41

AGCGGGTCTCGGTTGGGTTCCGCTAAATTTCTGTCCTGAGGCGGTGAGACTGAGTTTCATAGGGTCTCTGGGTCCCCGA
 ACCAGGAAGGGTTGAGGACACACATCTGCAAGCCCCGCCACCAAGTGAGGGCCCCGCTTTGGGGTCTCTCC
 TCCCTTTGCATTTCCACCCCTCCGGGCTTTGGGTCTTCTCTGGGGGACCCCTCCGCCGGAG**ATG**CGCGCGTTGATG
 CGGAGCAAGGATTCGTCTCTGCTGCCTGCTCTACTGGCCCGGGTGCTGATGGTGGAGAGCTCACAGATCGGCAGT
 TCGCGGCCAAACTCAACTCCATCAAGTCTCTCTGGCGGGGAGCGCTGTGTCAGGCCCGCAATCGATCTCCG
 GGCATGTACCAAGGACTGGCATTTCGGCGGCAGTAAGAAGGGCAAAACCTGGGGCAGGCCCTACCTCTTGATGAGT
 GATAAGGAGTGTGAAGTTGGGAGTTATGCCAGTCCCCACCAAGGATCATCGGCCTGCATGGTGTGTGTCGGAGA
 AAAAAGAGCGCTGCCACCAGATGGCATGTGCTGCCCCAGTACCCGCTGCAATATATGTCATCTGTATCCGAGTT
 ACTGAAAGCATTTTAACCCCTCACATCCCGCTCTGGATGGTACTCGGCACAGAGATCGAAACCAAGCTTACTTAC
 TCAAAACATGACTTGGGATGGCAAGATAGGAAAGCACACACTAAGATGTGCATATATAAAGGGCATGAAGGA
 GACCCCTGCCTACGATCATCAGACTGCATTTGAAGGGTTTGTCTGTCTCGTCATTCTGGACCAAAATCTGCAAA
 CCAGTGTCCATCAGGGGGAAGTCTGTACCAAAACACGAAGAAGGGTTCTCATGGGCTGGAATTTTCCAGCGT
 TGGCAGTGTGCGAAGGGCCTGTCTTGCAAAAGTATGAAAGATGCCACTACTCCTCCAAAGCCGAGACTCCATGTG
 TGTGAGAAAT**TGAT**CAACATTGAGGAACATCATCAATTGCAGACTGTGAAGTTGTGTATTTAATGCATTATAG
 CATGGTGGAAATAAGGTTTCAGATGCAGAAGAAATGGCTAAAATAAGAAACGTGATAAGAAATATAGATGATCACAA
 AAAGCGAAGAAAGAAACATGAATGAAATAGATAGAAATGGGTGACAAATGCAGTGCAGCCAGTGTTCATATTATG
 CAACTTGTCTATGTAAATAATGTACACATTTGTGGAATGCTATTATTAAGAGAACAAGCACACAGTGGAAAT
 ACTGATGAAGTAGCATTGCACTTTCCAAGAGTTTAGGTGTGCTGGAGGAGAGTTTCTTCAGATTGCTGATTGTC
 TTATACAAATTAACCTACATGCCAGATTTCTATTCAACGTTAGAGTTTAAACAAATACTCCTAGAATAACTTGTTA
 TACAAATAGGTTCTAAAATAAAATTTGCTTAACCAAGAAATGAAACATGGAGCAATTGTAAATTTACAACGAAAT
 TTACCTTTTGATTGTAACTACTCTCTCTTCAATCAAGAGTCTGGTAGATAAGAAAAAATCAGTCAATAT
 TTTCCAAATATTGCAAAATAAAGGCCAGTTCTTTAGGAAGGCCCTTAGGAAGACAATAAATAACAAACAAACAG
 CCACAAATACTTTTTTTTCAAAATTTTAGTTTTACTCTTAATTAAGAAGCTGATCAAGACAAAAACAGTTCC
 TTCAGATTCTACGGAATGACAGATATATCTCTCTTTATCCTATGTGATTCTCTGCTGAATGCATATATTTTCCA
 AACTATACCCATAAATGTGACTAGTAAATACTTACACAGAGCAAAATTTTCACATGAGCAACAAATTTTAA
 GATGTCCAATATATGTGGGAAAAAGAGCTAACAGAGAGATCAATTATTTCTTAAAGATTGGCCATAACCTATATTT
 GATAGAATTAGATTGGTAAATACATGATCATACATCTCTGTGGTAATAGACACTACACACATAGATCATATCTCTG
 CACTGGAGTAAGCAAGAAATTTGGGAAACCTTTTGTGTTTGTGAGGTTTGGCCAGTGGCAATATTTTAA
 AGGCACAAGTTGGCTGTTCACTTTTGAACACAGGGGATGCACAGTCTTAATGAATATCTGCATGGGATTTGCTAT
 CATAAATTTTACTATGCGAGATGAATTCAGTGTGAGGTCTGTGTCCGTACTATCCTCAAATTTATTTTATAG
 TGCTGAGATCCTCAAATAATCTCAATTTCAAGGAGTTTACAAAAATGACTCCTGAAGTAGACAGAGTAGGAGG
 TTTCAATTGCCCCCTATAAGCTTCTGACTAGCCAAATGGCATCATCCAATTTTCTGCCAAACCTCTGCGAGCATGT
 CTTTATTGCCAAAGGGCTAGTTTCGGTTTCTGCAGCCATTGGGGTTAAAAAATATAAGTAGGATAACTTGTAA
 ACCTGATATTGCTAATCTATAGACACCACAGTTTCTAAATTTCTTGAACACCTTTACTACTTTTTTTAAACCT
 AACTCAGTTCTAAATACTTTTCTGAGACCAAAAACATAAAGGTTATCTTATAGTCTGAGCTTTAAAGCTTTG
 TAGACACAATTCACCTTTTGTGTTTCTTTTACTTTAAATCCCATCTGCAGTCTCAAAATTTAAGTTCTCCAGTAG
 AGATTGAGTTTGAAGCTGTATATCTATTAAAAATTTCAACTTCCACATATATTACTAAGATGATTAAAGACTTA
 CATTTCCTGACAGGCTGCAAAAACAAAAATATAAACTAGTCCATCCAAGAACAAAGTTGTATATAACAGGT
 TGCTATAAGCTTGTGAATGAAAAATGGAACATTTCAATCAACATTTTCTATATAACAAATTTATTTATTTTACAAT
 TTGGTTTCTGCAATTTTCTTATGTCCACCTTTTAAAAATTTATTTTGAAGTAATTTTATTTACAGGAAATG
 TTAATGAGATGATTTTCTTATAGATATTTCTTACAGAAAGCTTTGTAGCAGAAATATATTGACGATTTGAC
 TTTGTAAATTTAGAAAAATGTATAATAAGATAAACTTATAAATTTTCTCCTCTAAAAACTGAAAAAATAA
 AAAAAAATAAATAAATAA

FIGURE 42

MAALMRSKDSSCCLLLLA AVL MVESSQIGSSRAKLNSIKSSIGGETFGQAANRSAGMYQGLA
FGGSKKGKNLGQAYPCSSDKECEVGRYCHSPHQGSSACMVCRRKKRCHRDGMCCPSTRCNN
GICIPVTESILTPHIPALDGT RHRDRNHGHYSNHD LGWQNLGRPHTKMSHIKGHEGDPCLRS
SDCIEGFCCARHFWT KICKPVLHQGEVCTKQRKKGSHGLEIFQRCDCAKGLSCKVWKDATYS
SKARLHVCQKI

Signal peptide:

amino acids 1-25

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

FIGURE 43

TCTCAATCTGCTGACCTCGTGATCCGCGCTGACCTTGTAATCCACCTACCTTGGCCTCCCAAA
 GTGTTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAACATCACGTTTTTAAAAATTGATT
 TCTTCAAATTCATGGCAAATATTTCCCTTCCCTTAACTTCTTATGTGAGAATGAGGAAGGA
 TAGCTGCATTTATTTAGTCAGTTTTTCATTGCATAGTAATATTTTCATGTAGTATTTTCTAAG
 TTATATTTTAGTAATTCATATGTTTTAGATTATAGGTTTTAACATACTTGTGAAAACTCTTG
ATGTGTTTTAAAGCCTTGGGCAGAATTCGTATTGTTGAGGATTGTGTTCTTTTATCCCCCT
 TTAAAGTCATCCGTCCTTGGCTCAGGATTGGAGAGCTTGCACCACCAAAAAATGGCAAACA
 TCACCAGCTCCCAGATTTTGGACCAAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACC
 CCAAGTACACAGAGAATAGTACAAGTCACCCTACAACCTACTACTTCTTGGGACCTCAAGCC
 CCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAAATCTCAACCTGAGCCATCCC
 CAGTTCCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCAGTGTCTCTCT
 CCTGGTTTGGAGTCTTTCTTCTCCAGGCAAACTTCGAGAATCAACACCTGGAGACAGTCC
 CTCCTACTGTGAACAAGCTTTTGAGCTTCCAGCAGCACCATTGAAAAATATCTCTGTGCTG
 TCCACCAGCCACAGCCCAACACATCAAACCTTGCTAAGCGGCGGATACCCCAAGTTCTAAG
 ATCCCAGCTTCTGCAGTGGAAATGCCTGGTTCAGCAGATGTCACAGGATTAAATGTGAGTT
 TGGGGCTCTGGAATTTGGGTGAGAACCTTCTCTCTGAATTGGATCAGTCCAAGCAGTG
 AAAATAGTAATCAGATTCCTATCAGCTTGATTGGAAGCTTTAAGTGAGCCTTTGAATACA
 TCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAACCTCCGTCATTACCTCCTG
 CAGTCTGACAAGCTCATCACTGAATTCGTAGTCCAGTAGCAATGTCTTCTCTTATGACC
 AGAGTTCTGTGCATAACAGGATCCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGTCCA
 GGAACCATCATGAATGGACATGGTGGTGGTGAAGTCAGCAGACACTAGACAGTAAGTATAG
 CAGCAAGCTACTCTGTGATGGCTGGTGCCAACCAACAGAGGAAGAGGATAGCTCAGGTGA
 TGTGGAACACACAGTTGGTCAATGGCTCATTTCGTTAAAAAGCAGCCCTTTTGCTTTTTTGT
 TTTTGGACCAAGTGTGGCTGTGGTGTTATTAGAAATGCTTAACCACAGCAAGAAGGAGGT
 GGTGGTCTCATATTTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATACAGTATGCAT
 TTTAAAGATGCTTGGGCCAGGCGGGGTGGCTGATGCCCATAAATCCAGTGCTTTGGGGGGCC
 AAGGCAGGCAGATTGCCAAGCTCAGGAGTTTGAAGCACCCTGGGCAACATGGTGAAACTC
 TGTCTTACTAAATACGAAAACTAGCCGGGTGTGGTGGCGCGCGTGCCTGTAATCCCAG
 CTACTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCC
 GTCTGAAAAGA

FIGURE 44

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTT
 PSTQQNSTSHPTTTTSWDLKPPTSQSSVLSHLDFKSQPEFSPVLSQLSQRQQHQSQAVTVPP
 PGLSEFSPQAKLRESTPGDSPSTVNKLLQLPSTTIENISVSVHQPPQPKHIKLAKRRIPPASK
 IPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAPSSSENSNQIPISLYSKSLSEPLNT
 SLSMTSAVQNSTYTTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPIYQSPVSSSESAP
 GTIMNGHGGGRSQQTLDISKYSSKLLLSWLVP TKQRKRIAHVMWKT PVGQWLIR

Signal peptide:

amino acids 1-24

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

FIGURE 45

GCCGAGTGGGACAAAGCCTGGGGCTGGGCGGGGGCC**ATGG**CGCTGCCATCCCGAATCCTGCT
 TTGGAAACTTGTGCTTCTGCAGAGCTCTGCTGTTCTCCTGCACCTCAGCGGTGGAGGAGACGG
 ACGCGGGGCTGTACACCTGCAACCTGCACCATCACTACTGCCACCTCTACGAGAGCCTGGCC
 GTCCGCTGGAGGTACCGACGGCCCCCGGCCACCCCGCCTACTGGGACGGCGAGAAGGA
 GGTGCTGGCGGTGGCGCGCGGCGCACCCGCGCTTCTGACCTGCGTGAACCGCGGGCACGTGT
 GGACCGACCGGCACGTGGAGGAGGCTCAACAGGTGGTGCACCTGGGACCGGCAGCCGCCCGGG
 GTCCCGCACGACCGCGCGGACCGCCTGCTGGACCTCTACGCGTCGGGCGAGCGCCGCGCCTA
 CGGGCCCCCTTTTCTGCGCGACCGCGTGGCTGTGGGCGCGGATGCCTTTGAGCGCGGTGACT
 TCTCACTGCGTATCGAGCCGCTGGAGGTGCGCGACGAGGGCACCTACTCCTGCCACCTGCAC
 CACCATTACTGTGGCCTGCACGAACGCCGCGCTCTTCCACCTGACGGTCGCGGAACCCACGC
 GGAGCCGCCCCCCCCGGGGCTCTCCGGGCAACGGCTCCAGCCACAGCGGCGCCCCAGGCCACG
 ACCCCACACTGGCGCGCGGCCACAACGTCAATGTCACTCGTCCCGGAGAGCCGAGCCAC
 TTCTTCCAGCAGCTGGGCTACGTGCTGGCCACGCTGCTGCTCTTCACTCCTGCTACTGGTCAC
 TGTCTCTCTGGCGCCCCGAGGCGCGCGGAGGCTACGAATACTCGGACCAAGAAGTCGGGAA
 AGTCAAAGGGGAAGGATGTTAACTTGGCGGAGTTCGCTGTGGCTGCAGGGGACCATGCTT
 TATACAGGAGTGAGGACATCCAGCTAGATTACAAAAACAACATCCTGAAGGAGAGGGCGGAGCT
 GGCCCAAGCCCCCTGCTGCCAAGTACATCGACCTAGACAAAGGGTCCGGAAGGAGAAT
 GCAAA**TAG**GGGAGGCCCTGGGCTCCTGGCTGGGCCAGCAGCTGCACCTCTCCTGTCTGTGCTC
 CTCGGGGCATCTCCTGATGCTCCGGGGCTCACCCCCCTTCCAGCGGCTGGTCCCGCTTTCTCT
 GGAATTTGGCCTGGGCGTATGCAGAGGCGCCTCCACACCCCTCCCCAGGGGCTTGGTGGC
 AGCATAGCCCCACCCCTGCGGCTTTGCTCACGGGTGGCCCTGCCACCCCTGGCACAAACC
 AAAATCCCCTGATGCCCATCATGCCCTCAGACCCCTTCTGGGCTCTGCCCGCTGGGGGCTTG
 AAGACATTCTGGAGGACACTCCCATCAGAACCTGGCAGCCCCAAAAGTGGGCTCAGCCTCA
 GGGCAGGAGTCCCACTCCTCCAGGGCTCTGCTCGTCCGGGGCTGGGAGATGTTCTTGAGGA
 GGACACTCCCATCAGAACTTGGCAGCCTTGAAGTTGGGGTCAAGCCTCGGCGAGGAGTCCCACT
 CCTCCTGGGGTGCTGCCTGCCACCAAGAGCTCCCCACCTGTACCACCATGTGGGACTCCAG
 GCACCATCTGTTCTCCCCAGGGACCTGCTGACTTGAATGCCAGCCCTTGCTCCTCTGTGTTG
 CTTTGGGCCACCTGGGGCTGCACCCCTGCCCTTTCTCTGCCCCATCCCTACCTAGCCTTG
 CTCTCAGCCACCTTGATAGTCACTGGGCTCCCTGTGACTTCTGACCTGACACCCCTCCCTT
 GGACTCTGCTGGGCTGGAGTCTAGGGCTGGGGCTACATTGGCTTCTGTACTGGGTGAGGA
 CAGGGGAGGAGTGAAGTTGGTTTGGGGTGGCCCTGTGTTGCCACTCTCAGCACCCACATTT
 GCATCTGCTGGTGGACCTGCCACCATCACATAAAGTCCCCATCTGATTTTTAAAAA
 AAAAA

FIGURE 46

```

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60618
<subunit 1 of 1, 341 aa, 1 stop
<MW: 38070, pI: 6.88, NX(S/T): 1
MALPSRILLWKLVLQSSAVLLHSAVEETDAGLYTCNLHHHYCHLYESLAVRLEVTDGPPAT
PAYWDGEKEVLAVARGAPALLTCVNRGHVWTD RHVEEAQQVVHWDRQPPGVPHDRADRLDL
YASGERRAYGPLFLRDRVAVGADAFERGFSLRIEPLVADEGTYSCHLHHHYCGLHERRVF
HLTVAEPHAEPPIRGSPGNGSSHSGAPGPDPTLARGHNVINVIVPESRAHFFQQLGYVLATL
LLFILLLVTVLLAARRRRGGYEYSDQKSGKSGKDVNLAEFAVAAGDQMLYRSEDIQLDYKN
NILKERAELAHSPLPAKYIDLKDGFRKENCK

```







































Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 237-262

N-glycosylation site.

amino acids 205-208

Cell attachment sequence.

amino acids 151-154

Coproporphyrinogen III oxidase proteins.

amino acids 115-140

FIGURE 47

CGCCGGAGGCAGCGGCGGCGTGGCGCAGCGGCGAC**ATG**GCCGTGTGCTCAGAGGACGACTTT
 CAGCACAGTTCAAACCTCCACCTACGGAACACAAAGCAGCAGTCTCCGAGCTGACCAGGAGGC
 ACTGCTTGAGAAGCTGCTGGACCGCCCGCCCTGGCCTGCAGAGGCCGAGGACCGTCTCT
 GTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCAGTCTACTGCCATGGAACCTCTTT
 ATCACTGCCAAGGAGTACTGGATGTTCAAACCTCCGCAACTCCTCCAGCCAGCCACCGGGGA
 GGACCCGTAGGGCTCAGACATCCTGAACCTACTTTGAGAGCTACCTTGCCGTGCTCCACCG
 TGCCCTCCATGCTGTGCCCTGGTGGCCAACTTCCTGCTTGTCACAGGGTTGCAGTCCACATC
 CGTGTCTGGCCTCACTGACGGTCATCCTGGCCATCTTCATGGTGATAACTGCCTGGTGAA
 GGTGGACACTTCCTCTGGACCCGTGGTTTTTTTGCGGTCACCAATTGCTGTCATGGTGATCC
 TCAGCGGTGCTCCACTGTCTTCAGCAGCAGCATCTACGGCATGACCGGCTCCTTTCCTATG
 AGGAACCCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCGTGGCCTC
 ATTGGTGGACTTGGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCTGACGG
 CCACCATCTTCTCGTGCTCTGCATGGGACTCTACCTGCTGCTGTCAGGCTGGAGTATGCC
 AGGTACTACATGAGGCCTGTTCTTGCGGCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCA
 GGACTCCCTCAGTGCCCTTCGGTGGCCTCCAGATTCATTGATCCACACACCCCTCTCC
 GCCCATCTCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTACCTACGTCTTCTTCATCACC
 AGCCTCATCTACCCCGCGCTGTCACCAACATCGAGTCCCTCAACAAGGGGCTCGGGCTCACT
 GTGGACCACCAAGTTTTTCATCCCCCTCACAACCTTCTCCTGTACAACCTTGTGACCTAT
 GTGGCCGGCAGCTCACCGCTGGATCCAGGTGCCAGGGCCCAACAGCAAGGCGCTCCAGGG
 TTCGTGCTCTCCGGACCTGCCTCATCCCCCTCTTCGTGCTCTGTAACCTACCAGCCCGCGT
 CCACCTGAAGACTGTGGTCTTCAGTCCGATGTGTACCCCGCACTCCTCAGCTCCCTGCTGG
 GGCTCAGCAACGGCTACCTCAGCACCTTGGCCCTCCTCTACGGGCTAAGATTGTGCCCAGG
 GAGCTGGCTGAGGCCACGGGAGTGGTGATGTCCTTTATGTGTGCTTGGGCTTAACACTGGG
 CTCAGCTGCTCTACCTCTCTGGTGCACCTCATCT**TAGA**AGGGAGGACACAAGGACATTGGTG
 CTTAGAGCCCTTTGAAGATGAGAAGAGAGTGCAAGAGGGCTGGGGGCCATGGAGGAAAGGCC
 TAAAGTTTCACTTTGGGGACAGAGAGCAGACACTCGGGCCTCATCCCTCCCAAGATGCCA
 GTGAGCCACGTCCATGCCCATTCGGTGCAAGGCAGATATTCAGTCATATTAACAGAACACT
 CCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACAGCTGATGGTTA
 ACATTCACACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTCGCCCTAGAGTTATTACA
 AAGCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCACTGCGCTCATTCAGCT
 GACAGCGAGATGCAAGCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGA
 AGTCCCTTGGCATGGTCAGTCTCAGGCCCAAGACTCAAGTGTGCACAGACCCCTGTGTTCT
 GCGGGTGAACAACCTGCCCACTAACCAGACTTGAAAAACCCAGAAAGATGGGCCCTTCATGAAT
 GCTTCATTCCAGAGGGACAGAGGGCCTCCTGTGCAGGGATCAAGCATGTCTGGCCTGGG
 TTTTCAAAAAAAGAGGATCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTCT
 TTTAGTGTTCTGTTTACAACATGTCAAAGCCATTGGTTCAAGGCGTAATAAATACTTGC
 GTATTCAAAA

FIGURE 48

MAVSEDDFQHSSNSTYGTSSSLRADQEALLEKLLDRPPGLQRPEDRFCGTYYIFFSLGI
 GSLLPWNFFITAKEYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFL
 LVNRVAVHIRVLASLTVILAI FMVITALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSI
 YGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSDVRNSALAFFLTATIFLVLCMGLY
 LLLSRLEYARYYMRPVLAHVFSGEEELPQDSL SAPSVASRFIDSHTFP LRPILKKTASLGF
 CVTYVFFITS LIYPVAVCTNIESLNKSGSLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVP
 GPNSKALPGFVLLR TCLIPLFVLCNYQPRVHLKTVVFQSDVYPALLSSLLGLSNGYLSTLAL
 LYGPKIVPRELAEATGVVMSFYVCLGLTLGSACSTLLVHLI

Transmembrane domain:

amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252,
 305-330, 448-472

FIGURE 49

GACAGTGGAGGGCAGTGGAGAGGACCGCGCTGTCCTGCTGTCAACCAAGAGCTGGAGACACCA
 TCTCCACCGAGAGTCA**ATG**GCCCCATTGGCCCTGCACCTCCTCGTCTCGTCCCCATCTCTC
 TCAGCCTGGTGGCCTCCAGGACTGGAAGGCTGAACGCAGCCAAGACCCCTTCGAGAAATGC
 ATGCAGGATCTGACTATGAGCAGCTGCTCAAGGTGGTGACCTGGGGGCTCAATCGGACCCT
 GAAGCCCCAGAGGGTGATTGTGGTTGGCGCTGGTGTGGCCGGGCTGGTGGCCGCCAAGGTGC
 TCAGCGATGCTGGACACAAGGTCACCATCTCGGAGGCAGATAACAGGATCGGGGGCCGCATC
 TTCACCTACCGGGACCAGAACACGGGCTGGATTGGGGAGCTGGGAGCCATGCGCATGCCAG
 CTCTCACAGGATCTCCACAAGCTCTGCCAGGGCTGGGGCTCAACCTGACCAAGTTCACCC
 AGTACGACAAGAACACGTGGACGGAGGTGCACGAAGTGAAGCTGCGCAACTATGTGGTGGAG
 AAGGTGCCCGAGAAGCTGGGCTACGCCTTGCCTCCCCAGGAAAAGGGCCACTCGCCCGAAGA
 CATCTACCAGATGGCTCTCAACCAGGCCCTCAAAGACCTCAAGGCACTGGGCTGCAGAAAGG
 CGATGAAGAAGTTTGAAAGGCACACGCTCTTGGAATATCTTCTCGGGGAGGGGAACCTGAGC
 CGGCCGGCCGTGCAGCTTCTGGAGACGTGATGTCCGAGGATGGCTTCTTCTATCTCAGCTT
 CGCCGAGGCCCTCCGGGCCACAGCTGCCTCAGCGACAGACTCCAGTACAGCCGCATCGTGG
 GTGGCTGGGACCTGCTGCCGCGCGCGCTGCTGAGCTCGCTGTCCGGGCTTGTGCTGTTGAAC
 GCGCCCGTGGTGGCGATGACCCAGGACCGCACGATGTGCACGTGCAGATCGAGACCTCTCC
 CCGCGCGCGGAATCTGAAGGTGCTGAAGGCGGACGTGGTGCTGCTGACGCGGAGCGGACCGG
 CGGTGAAGCGCATCACCTTCTCGCGCGCGCTGCCCCGCCACATGCAGGAGGCGCTGCGGAGG
 CTGCACTACGTGCCGGCCACCAAGGTGTTCTTAAGCTTCGCGAGGCCCTTCTGGCGGAGGA
 GCACATTGAAGGCGGCCACTCAAACACCGATCGCCCGTCCGCGCATGATTTCTACCCGCCGC
 CGCGCGAGGGCGCGCTGCTGCTGGCCTCGTACACGTGGTCGGACGCGGCGGCAGCGTTCCGC
 GGCTTGAGCCGGGAAGAGGCGTTGCGCTTGGCGCTCGACGACGTGGCGGCATTGCACGGGCC
 TGTCTGCGCCAGCTCTGGGACGGCACCGGCGTCTGCAAGCGTTGGCGGAGGACCAGCACA
 GCCAGGGTGGCTTTGTGGTACAGCGCGCGCGCTCTGGCAAACCGAAAAGGATGACTGGACG
 GTCCCTTATGGCCGCATCTACTTTGCCGCGGAGCACACCGCTACCCGCACGGCTGGGTGGA
 GACGCGGTCAAGTCCGCGCTGCGCGCGCCATCAAGATCAACAGCCGAAGGGGCCATGCAT
 CGGACACGGCCAGCCCCGAGGGGCACGCATCTGACATGGAGGGGCAGGGGCATGTGCATGGG
 GTGGCCAGCCAGCCCTCGCATGACCTGGCAAAGGAAGGACGAGCCACCTCCAGTCCAAGG
 CCAGTTATCTCTCCAAAACACGACCCACACGAGGACCTCGCAT**TAA**AGTATTTTCGGA
 AA

FIGURE 50

MAPLALHLLVLVPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVVVTWGLNRTLKPQRV
 IVVGAGVAGLVAAKVLS DAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRIL
 HKLCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGALRPQEKGHSPEDIYQMA
 LNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSEDGFFYLSFAEALR
 AHSCLSDRLQYSRIVGGWDL LPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPARNL
 KVLKADVLLLTASGPAVKRITFSPLPRHMQEALRRLHYVPATKVFLSFRPFVWREEHIEGG
 HSNTDRPSRMIFYPPPREGALLASYTWS DAAAAFAGLSREEALRLALDDVAALHGPVVRQL
 WDGTVVKRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWWETAVKS
 ALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQ
 NTTHTRTSH

Signal peptide:

amino acids 1-21

FIGURE 51

CTGACATGGCCTGACTCGGGAGACGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCGGGCCTTCTGCCTGCAT
 GACGCTCTGAAAGCCACCTGTCTCTGGAGGAACCCAGAGCGAGGGAGAAAGGACAGGGACTCGTGTGGCAGGAA
 GAACCTCAGAGCGGGGAGAGCCCACTTCTACTAGAAGCACTGAGAGATCGCGCCCCCTCGCAGGGCTCTGAATTTCCCT
 GCTGCTGTTTACAAAGATGCTTTTATCTTTAACTTTTGTGTTTCCCCACTCTCCGACCCCGGGCTTGATCTGCAT
 CCTGACATTTGGAGTGCCTATCTTTGTGGCTGATCACCAGACCTCAACCCGCTTTACTCTCTTCTTGACCTGAA
 CAATCAGTCTGTGGGAATTGAGGGAGGAGCAGCGAAGGGGGTTCCCAAGAAACAATGACCTTAACAAGTTGTCTG
 CTTCTCAGATGCCAAGACTATGTATGAGGTTTCCAAAGAGGACTCGCTGCTCTGCAATGGGCCCTGTGTTGGG
 ATATAGAAAACCAACAGCCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTCT
 CTTCTCTTGCATAAAGTTATAAATCATCACCAGACCAGTTTGTGGGCATCTTGTCTCAGAATAGGCCAGGTGAT
 GATCATCTCCGAATTGGCTTGTACACGTACTCTATGGTAGCTGTACCTCTGTATGACACCTTTGGGACCAGAGC
 CATCGTACATATTGTCAACAAGGCTGATATCGCCATGGTGATCTGTGACACACCCCCAAAGGCATTGGTGCTGAT
 AGGGAATGTAGAGAAAGGCTTCAACCCGAGCCTGAAGGTGATCATCTTATGGACCCCTTTGATGATGACCTGAA
 GCAAGAGGGGGAGAGAGTGGAAATTGAGATCTTATCCCTATATGATGCTGAGAACCCTAGGCAAGAGCACTTCAG
 AAAAAGCTGTGCCTCCTAGCCCAAGAACCTGAGCGTATCTGCTTACCAGTGGGACCAAGGTGACCCCAAGG
 AGCCATGATAACCCATCAAAATATTGTTCAAATGCTGCTGCCCTTCTCAAATGTGTGGAGCATGCTTATGAGCC
 CACTCCTGATGATGTGGCCATATCTACCTCCCTCTGGCTCATATGTTTGAAGGATTGTACAGGCTGTTGTGTA
 CAGCTGTGGAGCCAGAGTTGGATTCTTCCAAGGGATATTCCGTTGCTGGCTGACGACATGAAGACTTTGAAGCC
 CACATTTGTTTCCCGGGTGGCTCGACTCCTTAAAGGATCTACGATGAAGTCAAAATGAGGCCACAGCACCCTTT
 GAAGAAGTTCTTGTGTAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAAGGCTATCATCAGGCATGATAG
 TTTCTGGGACAAGCTCATCTTTGCAAGATTCAGGACAGCCTTGGCGGGAAGGTTCTGTGTAAITGTACACTGGAGC
 TGCCCCATGTCCACTTCAGTCATGACACTTCTCCGGGACGCAATGGGATGCTCAGGTTATGAAGCTTATGGTCA
 AACAGANTGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTCAATGAAGGAGAGGTCTGCAT
 TTTGCAATTACCTGAAGCTGGAAGTATGCTGTGCTGACATGAATCTTACAGTGAATGAAGGAGAGGTCTGCAT
 CAAGGGTACAAAGCTCTTCAAAAGATACCTGAAGGACCTGAGAGACCTGAGAAGACACAGGAGACCTTGAACGCTGATGGCTG
 GCTCTCACACAGGAGACATTGGTGGCTGGCTCCCGAATGGAAGCTCTGAGAGTATCGACCGTAAAAAGAACATTTT
 CAAGCTGCCCAAGGAGAGATCTTGCACAGAGAAGATAGAAAATATACAAACAGGAGTCAACCGATGTTACAT
 AATTTTTGTACACGGGAGAGCTTACGGTCTCTTAGTAGGAGTGGTGGTTCTTGACACAGATGTACTTCCCTC
 ATTTTCAGCCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGAACCTGTGCCAAAACCAAGTTGTAAGGGAGCCATTTT
 AGAAGACTTGCAGAAAATTTGGGAAGAAAGTGGCCTTAAACTTTTGAACAGGTCAAGGCCATTTTTCTTCATCC
 AGAGCCATTTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAAATACTT
 TCGGACCCAAATTGACAGCCTGTATGAGCAGATCCAGGATAGGATAAGGTACTTTAAGTACCTGCCGCGCACTG
 TGCATCTGCTTGTGAGAAAATGGATTAAAAACTATTCTTACATTTGTTTTGCCCTTCTCCTATTTTTTTTTAAACC
 GTTTAACTCTAAAGCCATAGCTTTGTTTTATATTAGACATATATATGTGTAAACTTAGTCTCCCAATTAATCA
 ATCCTGTCTTTCCCATCTTCCGATTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCTGTCTCTCAA
 GATCCAGTTTATGTTCTGTGCTCTTCTCATGATTTCCAACTTAATACTATTAGTAACCAAGTTCTCAAGGT
 CAAAGGGACCCCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAACTTGCCAAAGCTCTCATGCTTATTATCA
 TCTTCTACTGTTCAAACTAAGAGATTTTAAATTTCTGAAAACCTGCTTACAATTCATGTTTTCTAGCCACTCCAC
 AAAACCAATAAATTTTGTGTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTT
 CTGCGTAAATTAATTTGTGTACTGAAGGGAAAAAGTTGATCATACCAAACTTTCTTAACTCTCTAGTTAGATA
 TCTGACTTGGGATATAAATTTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTTCTTAAAGCATTATTTA
 CAGTAGGAAGCTGGGAGTAAATCTGTTCCCTACAGTTTGTGCTGAGCTGGAAGCTGTGGGGGAAGGAGTTGACA
 GGTGGCCCAAGTGAATTTTCCAGTAAATGAAGCAAGCACTGAATAAAACCTCCTGAACCTGGGACAAAGATCT
 ACAGGCAAGCAAGTGCACACACACAGGCTTATTTTCTGTGAAGGAACCACTGATCTCCCCACCCTTGGATT
 AGAGTCTCCTGCTACCTTACCCACAGATACACATGTTGTTTCTACTGTAAATGTAAGTCTTTAAATTAAC
 TATTACAGATAAAAA

FIGURE 52

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60775

<subunit 1 of 1, 739 aa, 1 stop

<MW: 82263, pI: 7.55, NX(S/T): 3

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFLLLFTKMLFIFN
FLFSPLPTPALICILTFGAAILFLWLITRPQPVLPLLDLNNQSVGIEGGARKGVSQKNNDLTS
CCFSDAKTMYEVEVFORGLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSS
PDQFVGIFAQNRPEWIISELACYTYSMVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKAL
VLIGNVEKGFTPSLVIIILMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPPSPED
LSVICFTSGTTGDPKAMITHQNIIVSNAAFLKCEHAYEPTPDDVAISYLPLAHMFERIVQ
AVVYSCGARVGFFQGDIRLLADDMKTLKPTLFPVPRLLNRIYDKVQNEAKTPLKKFLLKLA
VSSKFELQKGIIRHDSFWDKLIFAKIQDSLGGVRVIVTGAAPMSTSVMTFFRAAMGCQVY
EAYGQTECTGGCTFTLPGDWTSGHVGVPLACNYVKLEADVADMNYYFTVNNEGEVCIKGTNVFK
GYLKDPEKTQEALDSGDWLHTGDIGRWLPNGTLKIIDRKKNIFKLAQGEYIAPEKIENIYNR
SQPVLLQIFVHGESLRSSLVGVVVPDTPVLP SFAAKLGVKGSFEELCQNQVVREAILEDLQKI
GKESGLKTFEQVKAI FLHPEPFSIENGLLTPTLKAKRGELS KYFRTQIDSLYEHIQD

Important features:

Type II transmembrane domain:

amino acids 61-80

Putative AMP-binding domain signature.

amino acids 314-325

N-glycosylation site.

amino acids 102-105, 588-591 and 619-622

FIGURE 53

GGAGGCGGAGGCCGCGGCCGAGCCGGGCCGAGCAGTGAGGGCCCTAGCGGGGCCGAGCGGGG
 CCGGGGCCCTTAAGCCATTCTTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAAT
 CCGGT**ATGG**ACGACTGGAAGCCAGCCCTCATCAAGCCCTTGGGGCTCGGAAGAAGCGG
 AGCTGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGGCCCTGCGGAGATTCTGTCA
 GACAGGGGCCGTGCTTTTCTGCTGGTGACTGTCATTGTCAATATCAAGTTGATCTGGACA
 CTCGGCGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGCAAGACTATGATGAGGCC
 CTAGGCCGCTGGAGCCCCACGGCGCAGAGGCAGTGGTCCCCGGCGGGTCTTGACGTAGA
 GGTGATTCAAGTCGACGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGCTGGAGGATG
 AGGCCCGGGAGCAGGGCCGGGGCATCCATGTCATTGTCTCAACCAGGCCACGGGCCACGTG
 ATGGCAAAACGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGCTATTCTCT
 CAACATGGTAGCGCCCGCCGAGTGCTCATCTGCAGTGTCAAGGATGAGGGCTCCTTCCACC
 TCAAGGACACAGCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGC
 TGGAGGGACACATGGGCCCTCGTGGGACGAAAAGGAGGTCTGTCTTGGGGAGAAACATTC
 TAAGTCACCTGCCCTCTCTTCTGGGGGACCCAGTCTCTGCTGAAGACAGATGTGCCATTGA
 GCTCAGCAGAAGAGGCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGC
 TTCTGCAGCAAAGTTGAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCACACCCATCGA
 GTTCAGCCCTGACCCATCCAGACACAAGGTCTCAATGTGCTGTGGCTGTGCTGCTGCTGCTG
 GGAACCGACCAATACCTGTACAGGATGCTGCGCTCTGCTGCTTACAGCCAGGGGGTGTCT
 CCTCAGATGATAACAGTTTTCATTGACGGCTACTATGAGCAACCATGGATGTGGTGGCACT
 GTTTGGTCTGAGGGGCATCCAGCATACTCCCATCAGCATCAAGAATGCCCGCTGTGCTCAGC
 ACTACAAGGCCAGCCTCACTGCCACTTTCAACCTGTTTCCGAGGCCAAGTTGCTGTGGTT
 CTGGAAGAGGACCTGGACATTGCTGTGGATTTTTTCAGTTTCTGAGCCAAATCCATCCACCT
 ACTGGAGGAGGATGACAGCCTGTACTGCATCTCTGCCGTGGAATGACCAGGGGTATGAACACA
 TCGGCTGAGGACCCAGCACTGTACCGTGTGGAGACCATGCCCTGGGCTGGGCTGGGTGCTC
 AGGAGGTCTTGTACAAGGAGGAGCTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCGGGA
 TTGGGACATGTGGATGCGGATGCCGTGAACAACGCCGGGGCCGAGAGTGCAATCATCCCTGACG
 TTTCCCGATCTTACCACCTTTGGCATCGTCGGCCTCAACATGAATGGCTACTTTACAGGGCC
 TACTTCAAGAAGCACAAGTTCAACACGGTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCT
 GAAGAAAGAAGCTTATGAAGTGAAGTTTACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACC
 ACAGCAAGAACCCTTGTGAAGACTCTTTCTGCCAGACACAGAGGGCCACACCTACGTGGCC
 TTTATTCGAATGGAGAAAGATGATGACTTCAACACCTGGACCCAGCTTGCCAAGTGCTCCA
 TATCTGGGACCTGGATGTGCGTGGCAACCATCGGGGCTGTGGAGATTGTTTCGGAAGAAGA
 ACCACTTCCTGGTGGTGGGGTCCCCGCTTCCCCCTACTCAGTGAAGAAGCCACCTCAGTC
 ACCCAATTTTCTGGAGCCACCCCAAAGGAGGAGGGAGCCCCAGGAGCCCCAGAACAGAC
ATGAGACCTTCTCCAGGACCTCGGGGCTGGGTACTGTGTACCCCAAGCTTGCTGAGCCCT
 TCCCTCCATCTGTAGGATTTTGTAGATGCTGGTAGGGCTGGGGTACCTTGTTTTTAAACA
 TGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCCCCTGCTCCAACACCCCGTTCTGTAGTT
 AAAAGCTATTATTTACTTCTCTTGTGGAGAAGGGCAGGAGTACCTGGGAATCATTAGT
 ATCCCTAGCAGCTCATCTGCCCTTGAATACCTCACTTTCCAGGCCTGGCTCAGAACTTA
 ACCTATTATTGCTGTCTGTGAGGGCCTTGAACACAGGCCAACCTGGAGGGCTGGATTTC
 TTTTGGGCTGGAATGCTGCCCTGAGGTTGGGGCTGGCTTACTCAGGAACCTGCTGTGGC
 CAACCCATGGACAGGCCACGCTGGGGCCACATGCTGACACAGACTCACTCAGAGACCCCTTA
 GACATGGACAGGCCCTCTCTCAGCCCTCTCTTGTGCCAGATTTCAAAGCTGGATTAAAGTT
 GGTCAATGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 54

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA61185
><subunit 1 of 1, 660 aa, 1 stop
><MW: 75220, pI: 6.76, NX(S/T): 0
MDDWKPSFLIKFFGARKKRSWYLTWKYKLTNQALRRFCQTGAVLFLLVTIVIVNIKLILDTR
RAISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVL DVEVYSSRSKVYVAVDGTTVLEDEA
REQGRGIHVIVLNQATGHVMAKR VFD TYSPHEDEAMVLF LNMVAPGRVLI CTVKDEGSFHLK
DTAKALLRSLGSQAGPALGWRD TWAFVGRKGGPVFGEKHSKSPALSSWGD PVLLKTDVPLSS
AEEAECHWADTELNR RRRR RFC SKVEGYGSVCCKDPTPIEFSPDPLDNKVLNVPVAVIAGN
RPNYLYRMLRSLLSAQGVSPQMITVFIDGYEPEMDVVALFGLRGIQHTPISIKNARVSQHY
KASLTATFNLFPEAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSL YCISAWN DQGYEHTA
EDPALLYRVETMPGLGW LRRSLYKEELEPKWPTPEKLWDWDMWMMRMEQRRRGRECIIPDVS
RSYHF GIVGLNMNGYFHEAYFKKKHFNTVPGVQLRNVD SLKEAYEVEVHRL LSEAEVLDHS
KNPCEDSFLPDTEGHTYVAFIRMEKDDDFTTWTQLAKCLHIWDL DVRGNHRGLWRLFRKKNH
FLVVGVPASPYSVKKPPSVTPITFLEPPPKEEGAPGAPEQT
```

Important features of the protein:

Transmembrane domain:

amino acids 38-55

Homologous region to Mouse GNT1

amino acids 229-660

FIGURE 55

CGGACGCGTGGGCTGCTGGTGGGAAGGCCATAAGAACTGGAAAAGCCACTCTCTTGGAAACCACCACAC
 CTGTTTAAAGAACCTTAAGCACCATTAAAGGCCACTGGAAATTTGTTGTCTAGTGGTTGTGGGTGATA
 AAGGAGGGCAGAAATGGATGATTTTCATCTCCATTAGCCTGCTGCTCTCGGCTAGTTGTTGGTGGGATGTTA
 CGTGGCCCGGAATCATTCCCTTGGCTGTTAATTTTCAGAGGAACGACTGAAGCTGGTGACTGTTTTGG
 GTGCTGGCCTTCTCTGTGGAAGTCTCTGGGAGTCATCGTGGCTGAAGGAGTACATGCCCTTTATGAA
 GATATTTCTTGAAGGAAAAACCCCAAGCAAGTGAACACATAATGTGATTGCATCAGACAAAGCAGC
 AGAAAAATCAGTTGTCATGAACATGAGCAGCCACGACCACACACAGCTGCATGCCATATATTTGGTG
 TTTCCCTCGTTCTGGGCTTCGTTTTCATGTTGCTGGTGGACAGATTGGTAACCTCCCATGTGCATCTCT
 ACTGACGATCCAGAAGCAGCAAGGCTTAGCAATTCAAAATCACCACCAGCGTGGGTCTGGTTGTCCA
 TGCTGCAGCTGATGGTGTTCGTTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTTAATTG
 TGTTTTGGCAATCATGTACATAAGGCACCAGCTGCTTTTGGACTGGTTTCCCTTCTTGATGCATGCT
 GGCTTAGAGCGGAATCGAATCAGAAAGCACTTGCTGGTCTTTGCATTGGCAGCACCAGTTATGTCAT
 GGTGACATACCTAGGACTGAGTAAGAGCAGTAAGAAGCCCTTTAGAGGTGAACCCACGGGAGTGG
 CCATGCTTTTCTCGCCGGGACATTTCTTTATGTTGCCACAGTACATGTCTCCCTGAGGTGGGCGGA
 ATAGGGCACAGCCCAAGCCCGATGCCACGGGAGGAGAGGCCCTCAGCCGCTTGGAAATGGCAGCCCT
 GGTTCGTTGGTTCGCTCATCCCTCTCATCTGCTAGTAGGACACCAGCATTAATATGTTCAAGTCCAGC
 CTCTGTCAGGCGCGTTTGGCCATCCAGTGAGAACAGCCGCGTACAGCTACTCACTTCTCTCAGTC
 TCTTGTCTCACTTCGCGCATCTACATGTATTCTAGAGTCAGAGGGGAGGTGAGGTTAAAAACCTG
 AGTAATGGAAGAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTATAGACATCCCATTGTGT
 TATCTTTTAAAAAGGCCCTTGACATTTTGGCTTTTAAATATTTCTCTTAAACCTTATCTCAGGGAAGATG
 GAATTTAGTTTAAAGGAAAGAGGAGAACTTCATACTCACAATGAATAGTGATTATGAAATACAGT
 GTTCGTAAATTAAGCTATGTCTCTTCTTCTTAGTTTAGAGGCTAGCTGATTTATCCATTGATTTT
 AACATGGTTCCACCATTGTAAGACTGGTGCTTTAGCATCTATGCCACATGCGTTGATGGAAGGTGATA
 GCACCCACTCACTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAAATGATAGC
 AAGACACATGAAAGCTCTCTTTTACTCAAAAGAGATATCCATTGAAAGGGATGTCTAGAGGGAT
 TAAACAGCTCCTTTGGCAGCGTGCCTCTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGG
 TGGGAGGAGCTCTAAAGAGGTGACTGGTATTTTGTAGCATTCCTTGCAAGTTCTCCTTTGCAAGAT
 ACCTGTCTCCACATTCCTAGAGAGGAGCCAGTTCTAGTAGTTTCAGTTCTAGGCTTTCCTTCAAGAA
 CAGTCAGATCACAAAGTGCTTTTGGAAATTAAGGGATATTAATTTTAAAGTATTTTGGATGGTAT
 TGATATCTTTGTAGTAGCTTTTTTAAAGACTACCAAATGTATGGTTGCTCTTTTTTTTGTTTTT
 TTTTTTTTTAATTAATTTCTCTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGGTGAGCTTT
 GCGCAGACTGTGCTCTTCACATAAACACCTGTAGCAAGATGGATCATAAATGAGAAGTGTTTGCCTA
 TGATTTTAAAGCTTATTTGGAATCATGTCTCTTGTCTCTTCGTCCTTTCTTGTCTTTCTTCTAATCTT
 TCCCTAGCCTCTCCTCGCCACAATTGCTGCTTACTGCTGGTGTTAATATTTGTGGGATGAATT
 CTTATCAGGACAACCTCTCTCGAAGTGAATAATGAAGATAATAATATCTTTATCTTTATCCCTT
 CAAAGAAATACCTTGTGTCAAATGCCGCTTTGTTGAGCCCTTAAAGTACCACTCTCTATGTGTAA
 ATTGACACAATCATACTCTGGTAATTTAAACAATTTAGATAGCAAAAGTGTTTAAACAGACTAGGATA
 ATTTTTTTTTCAATTTTGGCAAAATTTTGTAAACCCCTGCTCTGTCAAATAAGTGTAATATTTGAT
 TATTAATTTATTTTACTTCTTACATACCATTTCAAAACACATTACACTAAGGGGGAACCAAGACTAGTT
 TCTTCAGGCGACTGGACGTAGTAGTTTGTAAAAACGTTTCTATGACGCATTAAGCTAGCATGCCATTT
 ATTTATTTTCTTATGTTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTTTGTGAGCCCTCTGCT
 GGCCACAGTGAGGAAAGTAGCACAAATAGGATACAGTTGTATGTAGTCAATTGGCAACAAATTCATACA
 ATTTTATCCAAGAGAAGGTATAGTATGGAAGTCCAATGACTTCTTGTATGGATGTTTAAAGCT
 GACTGGTGTAGACTTGAGGTTTCATCTAGTCCCTCAAACATATATGTTGCTAGATTCTCTCTGGA
 AACTGACTTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAA

FIGURE 56

MDDFISISLLSLAMLVGCYVAGIIPLAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHAL
 YEDILEGKHHQASETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQ
 IGNSHVHSTDDPEAARSSNSKITTTLGLVVHAAADGVALGAAASTSQTSVQLIVFVAIMLHK
 APAAFGLVSFLMHAGLERNRIRKHLVFAAAPVMSMVTYLGLSKSSKEALSEVNATGVAML
 FSAGTFLYVATVHVLPEVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIPILISVGHQH

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 37-56, 106-122, 211-230, 240-260, 288-304

FIGURE 57

GCTCGAGGCCGGCGGGCGGGAGAGCGACCCGGGCGGCTCGTAGCGGGGCCCGGATCCC
 CGAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAG**ATG**ATGGGCTTGG
 GAAACGGGCGTCGCAGCATGAAGTCGCCGCCCTCGTGCTGGCCGCCCTGGTGGCTGCATC
 ATCGTCTTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGAT
 CATGGAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGCGCCGTGGAGCTGAAGA
 AGAACGAGTTCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGC
 CACAACCTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGCGGTTTTTGGTGAA
 TAACATCACCCAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGA
 GGAATTACGGCAGGCTGCAGCAGGATGTCTCCAGTTTCAGAGAACCAGACCAACCTGGAG
 AGGAAGTTCTCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTG
 TGAGGAGCGAATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCAGAGACCTGA
 GTGAAAACAACGACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAG
 GCAGCAGGCTTGCCACACACAGAGGTGCCACAAGGAAGGGAACGTTGCTTGGTAACAGCAA
 GTCCCAGACACCAGCCCCAGTTCGGAAGTGGTTTTGGATTCAAAGAGACAAGTTGAGAAAG
 AGGAAACCAATGAGATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAG
 CCAGGCCGGGAGCAGGTGGTGAAGACAGACCTGTAGGTGAAGAGGCTTCGGGGGAGCCGG
 AGAACTGGGCCAGACCCACAGGTGCAGGCTGCCCTGTAGTGAGCCAGGAAAATCCAGAGA
 GTGGAGGGCCCTGAGCGAGACCAGCTTGTTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCT
 GCCGGGAAGGGAGAAACCAGCAGAACTGAGAGGAGAAGATGACTACAACATGGATGAAAA
 TGAAGCAGAATCTGAGACAGACAAGCAGCCCTGGCAGGGAATGACAGAAACATAGATG
 TTTTAAATGTTGAAGATCAGAAAAGAGACACCATAAATTTACTTTGATCAGCGTGAAAAGCGG
 AATCATACACTCT**TGA**ATTGAACCTGGAATCACATATTTCAACAGGGCCGAAGAGATGACTA
 TAAATGTTTCATGAGGGACTGAATACTGAAAACCTGTGAAATGTACTAAATAAATGTACATCTGA

FIGURE 58

MMGLGNGRRSMKSPPLVLAALVACIIVLGFNYWIASRSVDLQTRIMELEGRVRRAAAERGA
 VELKKNEFQGELEKQREQLDKIQSSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQL
 KTLQRNYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVA
 SRDLSENNDQRQQLQALSEPQPRLQAAGLPHTEVPOGKGNVLGNSKSQTPAPSSEVVLDSCR
 QVEKEETNEIQVVNEEPQRDRLPQEPGREQVVEDRPFVGGRGFGGAGELGQTPQVQAALSVSQ
 ENPEMEGPERDQLVIPDGQEEEQEAAGEGRNQQLRGEDDYNMDENEAESETDKQAALAGND
 RNIDVFNVEDQKRDTINLLDQREKRNHTL

Signal peptide:

amino acids 1-29



FIGURE 59

CGATGCGACAAAGGCTCAGTCTTGCTCTCTCGGCTGGGTCTGCTTCTCTCTACGCTGGCATTGGCCCTCTTCA
CAGTGGCTTCTGCTCCACCGCTTTGGAGCTACCAACAGTCAGACGCTCCAGAGGCCCGAGCCCTGGGTGCGT
TGACATCGGGGAGAGGAGGAAACCTGGGGCTCTGCTGATGGCTCCGATTTTCGGCGGGTGTGTGTTCGCTGTA
TGATGCTCTGCGATTTTACTTCCGCCACCCAGCATTCACAGTGGCTAGAGAGCTCTCTGTTCCTACCTCT
TCTTGGGCAAACTAAGCTTGTGACAGAGGATCTGAGAGTACAGCCGCCACATGGCCGGCTACAGCATCTCAGG
TTGACCTCTACACACACATCGACGGCTCTAAGGCGCTCAGCATGGCTCTACTGCTACCTTTATTGTATGCTGT
GTAGTAAATTTCGGCAGCCAGCGCTAGTGAAGAACATCTCATTAAAGCACTGCACAGTGCAGAGAGGCGTGTAG
TCTTCTACGGAGATGTATCTGGAAAGCACTTTCCCTGGTGCTTCTCCAAAGCTTTCTTCTTCCCTCTCTCT
ATTTACAGACCTAGACAGTGGACATGGCATGGCATCTGGAAACCTATCCACCACATGCAGACGTGGTGAATGC
AGCTGCTGATTTGCTCACTCTTGGGTGTGGACACTGTGGCCACAAGCATGGCCATCACCACTGGAAATGGCA
AGAACTTAGCCAGATGGACAGGTGATCAGGGACTTGTGGAGCTCTGGAGATGCACACACTCTGTGATGGTGT
CTGGGGACCATGGATGACCAAAATGGAGACCATGGAGGGAGCAGTGAAGCTGGAGTCTACTGCTCTCTTCTT
TGTAATGCCACAGAGCTTTCTCCGACAGCCACCCAGAGAGGAGCAGAGGTGATTTCTCAAGTCTCAAGCTTGTGC
CCACGCTGGCCCTGCTGCTGGGGCTGCCATCTGATTTGGGAATATCGGGAGATGTATGGCTGAGCTATCTTCAG
GGGTGAGGATCCCGCCGCCCTCTCTGTTTATGCCACAGCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT
CCCGATTTTCTTATCACTCAGCTCAGCTCTCAGCATCAGACTTCAGGATGGAGGCTTCATCAGCTGCAGAACTCT
TCTTCAAGGCTCTGCTGACTACAGTGGCTCTCAGAGGAGCCAAAGGGGGCTGAGGCGACTCGCATCTGTGA
TTGCTGAGCTGCAGAGTCTCTGGGGGGAGCTCGGGCACTGTGCATCAGTCTTGCGCTCGCTTCTCTCTGGTCC
GCAATGGCGGGGGTATGCTCTCTTGCTGCTCTCTGCTTATCTGCTGCTGGCATCTCAGTGGGCAATATCCC
CAGGCTTTCCATCTGCTCTCTACTCTGACACCTGTGGCTGGGGCTGTGGTGGGGCCATAGCCTATGCTGGGAC
TCTGGGAACTATTGAGCTGAAGCTAGATATAGTCTCTTAGGGGCTTGCGCTCAGTAGAGCTCATCTTCCCTTT
TCTGTGGAAGAGCTGGCTGCTCGGGGTCCAAAGAGGAGCTGGCAACCTGTTTCCATCTCTGGGCCGCTG
TGTTACTCTGCTGTTTCTTGCTGCTGTGTTCTTCTGTGATGTTAGTCTAGGCCAGGGCCACCCCTCT
TCTTTTGGGCTCTTCACTCTGCTCTGCTGTGTGTGCACCTTCTAGGGGAGGCGACTCTCCACCTAAGCTAC
CAACAATGCCCGGCTTTGGCATCTAGCCACAACAAACCCCAAGGCACATGTGTCATCTGCTTAGGCTTGT
GAATTGGTGTGTTTATGACAAAGGTACATGGGCTTTTTCATGTTGCCCTGAAGACACATGTTTGCCACT
CTCTCTCCGCTGAGTCTCTGTCATGCTACATGGTGGGTGTGCAAGCAAGAAATTTGATGATGAGCTGTGTGG
CGGCGCTGGTGGGCTGTGATGATGCCGTGCGCTTGTGCTCTGCCGCTCTGCTGATCTCAAGAGCCCGGAGCAC
CATGCTCTTTGTGGCTGGGAGTGCCTCAATGGCATGGGTATGCTGCTACTGGGCATTTGGCTCGGGGCT
CAGATGAGGCTCCCGGCTCTCGGGTCTGTGCTCTGGGAGCTCATGGTGTGCTGCTCGGGCTATGACAGGGC
TGCGCTCTAGGGCTCGCGTGTGCTTGAAGGCTGTGACAGTGGTGTGAGGCTGGGGAGGCGCTCAA
GGACAGCACTGTCTCACTCCCTTTTCAGGCCCCCCCACTCTTCAAGCTGACTTTGGATTATGTGTCCTCAAA
TCTACGACACATCAGGAGGAGTTTCGGGGCGGTTAGAGAGGACAAATCTCAGGGTCCCTGATCTGGCTGT
CTTATCACTTGGGAGTGTCTACTCTGCTCTGCTGTGTTCTGTGACAGACTTCTCTCTACATCTGCTGTGT
TGATCGGAGCGCATCAGCTCTTGTTCTGCTCTCTGTTCTTGTGACAGACTTCTCTCTACATCTGCTGTGT
CTGGATACCCGCTACCAACCCTGTGCTCTTTACTGTGCCATGGACGAGTCTCGGCTGGGCCCTATGGCA
CAGACACTCTACTCAACAGGCCACAGGCTCTTTTCAGCATCAATTGGCATCAGCACTTGTGGGATTC
CAGAGGCTCATGGCTCTGTACTTGGCTGCTGTTGCTGATGGGAGCAACAACTTTCGCTCCCACTCTCTCT
TTGCACTAGGTTGCCACTGTCTCTGCTCTGCTGCTTCTGTTGAGAGTCAAGGCTCGGGAAGACAGACAGC
CCCCGGAATGAAGCTGATCGGAGACTCAGCCAGGAGGGAAGAGGACCATGATGAGATCGCGCTCCGGG
ATGCGCTCAGCACTCTATGACAGACTGTGCGAGTGGGCTCAAGTCACTTTTATCTGGTATTCAAGTCT
TGCGCTGTGCTTGGACAGCTCTGCTCTCGACGATCTCATGCTGGAAGTGTGTTGCCCTAAGTTCATAT
TGGAGCTGGGGCTCAITGTGACAGCTGGGACTCTCTGGGCTCAAGTCTTGGTATGATGAGTGGTGTGTG
CTGAGCTCTCTGTGTCAGGCACTATTCTGGCCACAGAGGTGCTAGTCTGTGATTATCGCATCTGGCTTGGCT
ACAGAGAGTCTGGAGAAAGCTGTAGCTGGCTGTACAGTACTGGATGATTCGAAGACAGGCTCAGCCATCA
TCTTACTATCATGACGAGGGGGCGCTGACATCAGGACTCATTTATCTCAATTGAGGACAGTGGGATTA
TGCTCCCTAATCTGATTTGGATGATCATCTGAGGACAGAGGGGGCGTCTCCGAATGGAAATAAAATAGCGCG
GGTGGTGACTTGCACTATAATCCAGCACTTTGGGAGCAGAGTGGGAGATGCTTGGTCCAGGATTCAG
AGGAGGCTCTGGACATAACAAGGACCGTCTCATCTATTAAAAAAGGTGATAAAATGATAAT

FIGURE 60

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809
```

```
<subunit 1 of 1, 1089 aa, 1 stop
```

```
<MW: 118699, pI: 8.49, NX(S/T): 2
```

```
MQKASVLLFLAWVCFLFYAGIALFTSGFLLTRLELTNHSSCQEPFPGSLPWGSQKPGACW
MASRFSRVVLVIDALRFDAQPPQHSHVPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQ
VDPPTTTMQRLKALTTGSLPTFTDAGSNFASHAIVEDNLIKQLTSAGRVRVFMGDDTWKDLF
PGAFSKAFFFPFSFNVRDLDTVDNGILEHLYPTMDSGEWDVLIHFLGVDHCHGKHGPHHPHEM
AKKLSQMDQVIQGLVERLENDTLLVVAGDHGMTTNGDHGGDSELEVSAAFLYSPYAVFPST
PPEEPEVPIQVSLVPTLALLLGLPIPFNGNIGEVMAELFSGGEDSQPHSSSALAQAASALHNAQ
QVSRFLHTYSAATQDLOAKELHQLQNLFSKASADYQWLLQSPKGAEATLPTVIAELQQFLRG
ARAMCIESWARFSLVRMAGGTALLAASCFCICLLASQWATSPGFFPCPLLLTPVANGVLGAI
YAGLLGTIELKLDLVLGAVAAVSSFLPFLWKAWAGWGSKRPLATLFPPIPGVLLLLLFLRLA
VFFSDSFVVAEARATPFLGSGFILLVVLHVEGQLLPKLLTMPRLGTSATNPPRHNGAY
ALRLGIGLLLLTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVAALVALLA
AVRLWLRLRRYGNLKSPEPPMLFVRWGLPLMALGTAAYWALASGADEAPPRLRVLVSGASMVLP
RAVAGLAASGLALLLWKVPVTVLVKAGAGAPRTRTVLTPFSGPPTSQADLDYVVPQIYRHMQE
EFRGRLERLTKSQGPLTVAAYQLGSGVSYAAMVTALTLLAFPLLLHAERISLVFLLFLQSFLL
LLHLLAAGIPVTPPGFPTVPWQAVSAWALMATQTFYSTGHQPVPFAIHWHAAVFVGFPEGHGS
CTWLPALLVGANTFASHLLFAVGCPLLLLPFLCESQGLRKRQPPGNEADARVRPEEEEEEP
LMEMRLRDAPQHFYAALLQLGLKYLFLGIQILACALAASILRRHLMVWKVFAPKFIFEAVG
FIYSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR
```

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domains:

amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850,
1016-1034, 1052-1070

Leucine zipper pattern.

amino acids 843-864

N-glycosylation sites.

amino acids 37-40, 268-271

FIGURE 61

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGT
 GTCTCTGGTGGTTTGCCATAAACCTGCAAAACATCACCTTCTTATCCATCAACATGAAGA**ATGT**
 CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATT
 TCATCACAAATTGGCCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGT
 CCTGACAGCTCCAGAGAAGTGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAA
 TATACTCCAATCTGAAGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCC
 CAGTGTGTGACCAACCACACGCTGGTGTCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGT
 ACACGTGGAGTCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTG
 CCAGGACTTTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTG
 CCCATATCTATTACCGTGTTCCTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCA
 CGTTGGCAAGAGAAAACCCAGCAAATTTGATTTTGATTATGGAATGAATTTGACAAAA
 GATTCTTTGTGCTCTGCTGAAAAATCGTGATTAACCTTTATCACCTCAATATCTCGGATGAT
 TCTAAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAGCAGTGATGTATCCAGCCTTAA
 TGATCCTCAGCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAG
 GGTATGCTTCGCATTTGATGGAATTTTTTGTGACTCTGAAGAAAACCGGAAGGTACTTCT
 CTCACCCAGCAAGAGTCCCTCAGCAGAACATAACCCCGGATAAAACAGTCATTGAATATGA
 ATATGATGTGAGAACCCTGACATTTGTGCGGGGCTGAAGAGCAGGAGCTCAGTTGTCAGG
 AGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCC
 CAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCGCAGGAGCA
 CACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACACCCTGGTCGACTGGGATCCCC
 AAAGTGGCAGGCTGTGTATTCCTTCGTGTCCAGCTTCGACCAGGATTGAGAGGCTGCGAG
 CCTTCTGAGGGGATGGGCTCGGAGAGGAGGCTCTCTATCTAGACTCTATGAGGAGCCGGC
 TCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGT
 TATATGTGCAGATGAAAAC**TGAT**GGCAACACTTCCTTTTGCTTTTGTTCCTGTGCAAAAC
 AAGTGAGTCACCCCTTTGATCCAGCCATAAAGTACCTGGGATGAAGAAGTTTTTTCCAGT
 TTGTCACTGTCTGTGAGAACTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTG
 GTTCATGCATGTAGGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGT
 GTTCTATGCAGAGAAAGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTACAG
 TGGGTGT

FIGURE 62

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62815

<subunit 1 of 1, 442 aa, 1 stop

<MW: 49932, pI: 4.55, NX(S/T): 5

MSYNGLHQRVFKELKLLTLCSSISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQ
QIYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSKQ
CARTLKDQSSEFKAKIIIFYVLPISITVFLFSVMGYSIYRIHVKGKHPANLILYIGNEFD
KRFFVPAEKIVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKH
LGYASHLMEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSL
QEEVSTQGTLLSQAALAVLGPTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWD
PQTGRLCIPSLSSFDQDSEGCEPSEGDLGEEGLLSRLYEPA PDRPPGENETYLMQFMEEW
GLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 63

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCG
 TCTGCC**ATG**GGGCTCGGGTTGAGGGGCTGGGGACGTCTCTGCTGACTGTGGCCACCGCCCT
 GATGCTGCCCCGTGAAGCCCCCGCAGGCTCCTGGGGGGCCAGATCATCGGGGGCCACGAGG
 TGACCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGCCAACATCACTGCGGA
 GGCTTCTGTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGACCT
 CCGCACTGGCCTGGTGGTGTCTGGGCGCCACGTCTGAGTACTGCGGAGCCCCACCCAGCAGG
 TGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGAC
 ATCTGCCTGCTGCGGCTGAACGGCTCTGCTGTCTGGGCCCTGCAGTGGGGCTGCTGAGGCT
 GCCAGGGAGAAGGGCCAGGCCCCCCACAGCGGGGACACGGTGCCGGGTGGCTGGCTGGGGCT
 TCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCCAAGGTCCGAGTGCTGGAC
 CCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCACCCGAC
 TTGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGGCCCTGGTGTGCAGGA
 ACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCCGAC
 GTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGACGTGGTTCGGCGGACGAGTCC
 CCAGCCCGGCCCCCTGCCTGGGACCACAGGCCCCAGGAGAAGCCGCC**TGAG**CCACAACT
 TGTGCGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCGTGGCTGGGCCCCACGGG
 AAGCCTGATGTTCAAGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACATGCA
 AAGGGCAGAAGCAAACCCAGTAAATGTTAACTGACAAAAAAAAAAAAAAAAAAGAAA

FIGURE 64

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62845
><subunit 1 of 1, 283 aa, 1 stop
><MW: 30350, pI: 9.66, NX(S/T): 2
MGLGLRGWGRPLLTVATALMLPVKPPAGSWGAIIGGHEVTPHSRPFYMASVREFGQHHCGGF
LLRARWVVSAAHCFSHRDLRTGLVVLGAHVLSTAEPQQVFGIDALTHPDYHPMTHANDIC
LLRLNGSAVLGPAVGLLRLLPGRRARPPPTAGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPD
VCNSSWKGHLLTLMCLTRSGDSHRRGFCSADSGGPLVCRNRAHGLVSFSGLWCGDPKTPDVY
TQVSAFVAWIWDVVRSSPQPGPLPGTTRPPGEAA
```

Signal peptide:

amino acids 1-30

FIGURE 65

GAGCTACCCAGGCGGCTGGTGTGCAGCAAGCTCCGCGCCGACTCCGGACGCCTGACGCCTGA
 CGCCTGTCCCCGGCCCGG**CATG**AGCCGCTACCTGCTGCCGCTGTCGGCGCTGGGCACGGTAG
 CAGGCGCGCCGCTGCTGCTCAAGGACTATGTACCGGTGGGGCTTGCCCCAGCAAGGCCACC
 ATCCCTGGGAAGACGGTCATCGTGACGGGCGCCAAACAGGCATCGGGAAGCAGACCGCCTT
 GGAAC TGCCAGGAGAGGAGGCAACATCATCTGGCCTGCCGAGACATGGAGAAGTGTGAGG
 CGGCAGCAAAGGACATCCGCGGGGAGACCCTCAATCACCATGTCAACGCCCGGCACCTGGAC
 TTGGCTTCCCTCAAGTCTATCCGAGAGTTTGCAGCAAAGATCATTGAAGAGGAGGAGCGAGT
 GGACATTCTAATCAACAACGCGGGTGTGATGCGGTGCCCCCACTGGACCACCGAGGACGGCT
 TCGAGATGCAGTTTGGCGTTAAACCACTGGGTCACTTTCTCTTGACAAACTTGTGCTGGAC
 AAGCTGAAAGCCTCAGCCCTTTCGCGGATCATCAACCTCTCGTCCCCTGGCCCATGTTGCTGG
 GCACATAGACTTTGACGACTTGAAC TGGCAGACGAGGAAGTATAACACCAAAGCCGCCTACT
 GCCAGAGCAAGCTCGCCATCGTCTCTTACCAAGGAGCTGAGCCGGCGGCTGCAAGGCTCT
 GGTGTGACTGTCAACGCCCTGCACCCCGCGTGGCCAGGACAGAGCTGGGCAGACACACGGG
 CATCCATGGCTCCACCTTCTCCAGCACCACTCGGGCCCATCTTCTGGCTGCTGGTCAAGA
 GCCCGAGCTGGCCGCCAGCCAGCACATACCTGGCCGTGGCGGAGGAAGTGGCGGATGTT
 TCCGGAAGTACTTTCGATGGACTCAAACAGAAAGCCCCGGCCCCGAGGCTGAGGATGAGGA
 GGTGGCCCGGAGGCTTTGGGCTGAAAGTGCCCGCCTGGTGGGCTTAGAGGCTCCCTCTGTGA
 GGGAGCAGCCCTCCCCAG**TAA**CTCTGGAGCAGATTGAAAGCCAGGATGGCGCCTCCAG
 ACCGAGGACAGCTGTCCGCCATGCCCGCAGCTTCTTGACACTACCTGAGCCGGGAGACCCAG
 GACTGGCGGCCCGCATGCCCGCAGTAGGTTCTAGGGGGCGGTGCTGGCCGCACTGGACTGGC
 CTGCAGGTGAGCACTGCCCGGGCTCTGGCTGGTTCCGTCTGCTCTGCTGCCAGCAGGGGAG
 AGGGGCCATCTGATGCTTCCCCTGGGAATCTAAACTGGGAATGGCCGAGGAGGAAGGGGCTC
 TGTGCACCTTGCAAGGCCACGTCAAGAGAGCCAGCGGTGCCTGTGCGGGAGGGTTCCAAGGTGC
 TCCGTGAAGAGCATGGGCAAGTTGTCTGACACTTGGTGGATTCTTGGGTCCCTGTGGGACCT
 TGTGCATGCATGGTCTCTCTGAGCCTTGGTTTCTTCAAGCAGTGAGATGCTCAGAATAACTG
 CTGTCTCCCATGATGGTGTGGTACAGCGAGCTGTTGTCTGGCTATGGCATGGCTGTGCCGGG
 GGTGTTTGCTGAGGGCTTCTGTGCCAGAGCCAGCCAGAGAGCAGGTGCAGGTGTATCCC
 GAGTTCAGGCTCTGCACGGCATGGAGTGGGAACCCACCAGCTGCTGCTACAGGACCTGGGA
 TTGCCTGGGACTCCCACCTTCTATCAATTCTCATGGTAGTCCAAACTGCAGACTCTCAAAC
 TTGCTCATTT

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64842

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35932, pI: 8.45, NX(S/T): 1

MSRYLLPLSALGTVAGAAVLLKDYVTGGACPSKATIPGKTIVIVTGANTGIGKQTALELARRG
 GNIILACRDMEKCEAAAKDIRGETLNHHVNNARHLDLASLKSIREFAAKIIEEEERVDILINN
 AGVMRCPHWTTEDGFEMQFGVNHGHLFTNLLLDKLKASAPSRIINLSSLAHVAGHIDFDD
 LNWQTRKYNTKAAAYCQSKLAIVLFTKELSRRLQGSGVTVNALHPGVARTELGRHTGIHGSTF
 SSTTLGPFIWLLVKPELAAQPSTYLAVAEELADVSGKYFDGLKQKAPAPEAEDEEVARLW
 AESARLVGLEAPSVREQPLPR

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162

FIGURE 67

GAAGTTCGCGAGCGCTGGC**ATG**TGGTCTTGGGGCGCGGCTGGCGGCGCTGCTGGCGGTGCTG
GCGCTCGGGACAGGAGACCCAGAAAGGGCTGCGGGCTCGGGGCGACACGTTCTCGGCGCTGAC
CAGCGTGGCGCGCGCCCTGGCGCCCGAGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGTC
GCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCTACGACAAGGTACTTTCTTTGCAT
GAGGATTCAACAACCCCTGTGGCTAACCCCTGCTGCTGCAATTTACTCTCATCAAACGCCCTGCA
GTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGG
ATGGCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGG
GCCCTGATCGGGCTGCAGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCGAGGTGTCTT
TCAGAGAGTCACTGGCTCTGCCATCACTGACCTGTACAGCCCCAAACGGCTCTTTTCTCTCA
CAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCATGACATGGGGATTATTACCATGCC
ATTCCATGGCTGGAGGAGGTGTAGTCTCTTCCGAGGATCTTACGGAGAGTGGAAGACAGA
GGATGAGGCAAGCTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCGAGGAA
ATGTTTCTGTGCCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCAGATAATAAGAGG
TATGTCAGACCTGGGTCTCCAGCCCACCTCTCTACAGATCCCTAGCCTCTACTGTTCTCTAT
GAGACCAATTTCAAACGCCCTACCTGCTGCTCCAGCCCATCCGGAAGGAGGTATCCACCTGGA
GCCCTACATTGCTCTCTACCATGACTTCGTGAGTCACTGAGAGGCTCAGAAATATAGAGAAC
TTGCAAGAACCATGGCTACAGAGGTGAGTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAG
TACCGCATCAGCAAAAGTGCCTGGCTGAAGGACACTGTTGACCCAAAACCTGGTGACCCCTCAA
CCACCGCATTGCTGCCCCACAGGCCCTTGATGTCCGGCCTCCCTATGCAGAGTATCTGCAGG
TGGTGAACATATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATGCTACGTCAACCAAGC
AGCCCCCTCTACAGAAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCTC
UGGTGGAAGCTGGAGGAGCCAGCCCTTCATCTATGCCAACCTCAGCGTGCCCTGTGGTTAGGA
ATGCAGCACTGTTTTGGTGGAACTTGACAGGAGTGGTGAAGGGGACAGTGACACACTTCAT
GCTGGCTGTCTGTCTGTGGTGGGAGATAAGTGGGTGGCCAACAAGTGGATACATGAGTATGG
ACAGGAATTCCGAGACCCCTGCAGCTCCAGCCCTGAAGACTGAAGTCTGTTGGCAGAGAGAAGC
TGGTGGAGTCTCTGTGGCTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTAGGAGAGGAGAA
AGCAGAGCAGCCTCCTGGAAGAAGGCCCTGTGAGCTTTGTCTGTGCCCTCGCAATCAGAGGC
AAGGGAGAGGTTGTTACAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAA
GTCAGAGTAGGATGCACAGTACAAAGGAGGGGGGAGTGGAGGCCTGAGAGGGAAGTTTCTGG
AGTTCAGATACTCTCTGTTGGGAACAGGACATCTCAACAGTCTCAGGTTCCGATCAGTGGGTCT
TTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGGCAATGAGGACACCTGCAGGAG
GGGCTAGCCTGACTCCAGAACTTTAAGACTTTTCCCCACTGTCCTTCTGCTGCAGCCCAAG
CAGGGAGTGTCCCTCCCAAGCATATCCAGATGAGTGATACATTATATAAGGATTTTT
TTTAAGTTGAAAACAACTTTCTTTTCTTTTGTATGATGGTTTTTAAACAGTCATTAAAA
ATGTTTATAAATCAAAA

FIGURE 68

MGPGARLAALLAVLALGTGDPERAAARGDTFSALTSSVARALAPERRLGLLRRYLRGEEARL
 RDLTRFYDKVLSLHEDSTTPVANPLLAFTLIKRLQSDWRNVVHSLSEASINIRALKDGYEKVE
 QDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRTVGSAITDLYSPKRLFSLTGDDCFQ
 VGKVAYDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRAGNVSCALS
 LSREFLLYSPDNKRMAVNLYKIERLLAESPNHVVAEAVIQRPNIPHLQTRDITYEGLCQTLGS
 QPTLYQIPSLYCSYETNSNAYLLQPIRKEVIHLEFYIALYHDFVSDSEAQKIRELAEPWLQ
 RSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRPPYAEYLVVNYGIG
 GHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLLSSVEAGGATAFIYANLSVPVVRNAALFWW
 NLHRSGEGDSDTLHAGCPVLVGDKWVANKWIEHYGQEFRRPCSSSPED

Signal peptide:

amino acids 1-19

FIGURE 69

GAGATAGGGAGTCTGGGTTTAAAGTTCTGCTCCATCTCAGGAGCCCTGCTCCCACCCCTAG
 GAAGCCACCAGACTCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCC
 ACGAGCGTGGCTGAGGGACCGAGCCGAGAGCCCCGAGCCCCCGTAACCCGCGCGGGGAG
 CGCCAGGATCGCGCGGGGACTCGGAGCAGGTGCGCTACTGCGCGCGCTTCTCCTACCTC
 TGGCTCAAGTTTTCACTTATCATCTATTCCACCGTGTCTGGCTGATTGGGGCCCTGGTCTT
 GTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCCTTGAAAGTGCCTTCC
 TGGCTCCAGCCATCATCCTCATCTCCTGGGCGTCTGTATGTTTATGGTCTCCTTCATTGGT
 GTGCTGGCGTCCCTCCGTGACAACTGTACCTTCTCCAAGCATTCATGTACATCCTTGGGAT
 CTGCTCATCATGGAGCTCATTGGTGGCGTGGTGGCCTTGACCTTCCGGAACCAGACCATTG
 ACTTCTGAACGACAACATTCGAAGAGGAATTGAGAACTACTATGATGATCTGGACTTCAA
 AACATCATGGACTTTGTTTCAAAAAAGTTCAAGTGTGTGGCGGGGAGGACTACCGAGATTG
 GAGCAAGAATCAGTACCACGACTGCAGTGCCTTGGACCCCTGGCCTGTGGGGTGCCTTACA
 CCTGCTGCATCAGGAACACGACAGAAGTTGTCAACACCATGTGTGGCTACAAAATATCGAC
 AAGGAGCGTTTCAGTGTGAGGATGTATCTACGTGCGGGGCTGCACCAACGCCGTGATCAT
 CTGGTTCATGGACAACACACCATCATGGCGTGCATCCTCCTGGGCATCCTGCTTCCCCAGT
 TCCTGGGGGTGCTGCTGACGCTGCTGTACATCACCCGGGTGGAGGACATCATCATGGAGCAC
 TCTGTCACTGATGGGCTCCTGGGGCCCGGTGCCAAGCCAGCGTGGAGGCGGACGACACGGG
 ATGCTGCTTGTGCTACCCCAATTAGGGCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTC
 TGGGATAGCACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCTCTGCCACAC
 CTCAGTACTGACCAAAGCCAGGGCTGTGTGTCCTGTGTGTAGTCCCACGGCCTCTGCCTC
 CCCAGGGAGCAGAGCCTGGGCCTCCCCTAAGAGGCTTTCCCCGAGGCAGCTCTGGAATCTGT
 GCCACCTTGGGGCCTGGGGAACAAGGCCCTCCTTTCTCCAGGCCTGGGCTACAGGGGAGGGA
 GAGCCTGAGGCTCTGCTCAGGGCCCATTTTATCTCTGGCAGTGCCTTGGCGGTGGTATTCAA
 GGCAGTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCTCGGGGACAGGAGGAAGG
 GCATCTGGGGAAGGGCAGGAGGGAAGAGCTGTCCATGCAGCCACGCCCATGGCAGGTTGGC
 CTCTTCTCAGCCTCCAGGTGCCTTGAGCCCTCTTGCAAGGGCGGCTGCTTCTTTGAGCCTA
 GTTTTTTTTACGTGATTTTTGTAAACATTCATTTTTTGTACAGATAACAGGAGTTTCTGAC
 TAATCAAAGCTGGTATTTCCCGCATGTCTATTCTTGCCCTTCCCCAACACAGTTTGTAA
 TCAACAATAAAAAACATGTTTTGTTTTGTTTTTAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863

><subunit 1 of 1, 294 aa, 1 stop

><MW: 33211, pI: 5.35, NX(S/T): 3

MPRGDSEQVRYCARFSYLWLKFSLIISTVFWLIGALVLSVGIYAEVERQKYKTLES AFLAP
AIILILLGVVMFVSFIGVLASLRDNLVLLQAFMYILGICLIMELIGGVVALTFRNQ TIDFL
NDNIRRG IENYYDDLD FKNIMDFVQKKFKCCGGEDYRDWSKNQYHDCSAPG PLACGV PYTCC
IRNTTEVVNTMCGYKTIDKERFSVQDVIYVRGCTNAVIIWFMDNYTIMACILLGILLPQFLG
VLLTLLYITRVEDIIMEHSVTDGLLGPGAKPSVEAAGTGCCLCYPN

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 22-42, 57-85, 93-116, 230-257

FIGURE 71

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCATTGCTGCCCTCTGA
 CACCTGGGAAG**ATG**CGCCGGCCCGTGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACC
 TTGATCCAAGCCACCCTCAGTCCCACTGCAGTTCTCATCTCGGCCCAAAAGTCATCAAAGA
 AAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCTCGCAGCAGCTGCCGCTGC
 TCAGTGCCATGCGGGAAAAGCCAGCCGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACC
 GTCCTGAAGCACATCATCTGGCTGAAGGTCACTCACAGTAAACATCCTCCAGCTGCAGGTGAA
 GCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCA
 ACACGCCCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATC
 CGCATGGACACCAGTGCAAGTGGCCCCACCCGCTGGTCTCAGTGACTGTGCCACCAGCCA
 TGGGAGCCTGCGCATCCAACCTGCTGTATAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGC
 AGGTCAATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTG
 ATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCATTTC
 CCTCAGCATTTGACCGCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTC
 AGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAGGTGACCAAGTGGTTCAATAAC
 TCTGCAGCTTCCTTGACAATGCCACCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCA
 GGACGTGGTGAAGAGTGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTCATGGTCTGTTGG
 ACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAG
 GCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTT
 TTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGTGGAAGTGTTCCTCCCA
 GTGAAGCCCTCCGCCCTTTGTTCACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTAC
 ACCAAAGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGAT
 GAACTCTGGGATTGGCTGGTTCACCTGATGTTCTGAAAACATCATCACTGAGATCATCC
 ACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTG
 AAGGCCTTGGGATTCGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCC
 AGCCTCCTTGTGGAACCCAGCTCTCTGTCTCCAG**TGA**AGACTTGGATGGCAGCCATCAG
 GGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAAT
 CAATAACACTTGCTGTGAAAAA

FIGURE 72

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881

><subunit 1 of 1, 484 aa, 1 stop

><MW: 52468, pI: 7.14, NX(S/T): 3

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDNATSILQQPLLSAM
 REKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPL
 VKTIVEFHMTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMN
 LLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYL
 GAKLLDSQGKVKWFNNSAASLTMTPLDNIPFSLIVSQDVVKAAVAVALSPPEEFMVLLDSVL
 PESAHRLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEAL
 RPLFTLGLIEASSEAQFYTKGDQLILNLNNISSDRIQLMNSGIGWFPQDVVKNIITEIIHSIL
 LPNQNGKLRSGVPVSVLVKALGFEEAESSLTKDVLVTPASLWKPPSPVQS

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 73

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GAGCGAACATGGCAGCGCGTTTGGCGGTTTTGGTGTGTCTCTGTGACCATGGTGGTGGCGCTG
 CTCATCGTTTGCAGACGTTCCCTCAGCCCTCTGCCCAAAGAAAGGAGATGGTGTATCTGA
 AAAGGTTAGTCAGCTGATGGAATGGACTAACAAAAGACCTGTAATAAGAATGAATGGAGACA
 AGTTCCTCGTCGCCCTTGTGAAAGCCCCACCGAGAAATTACTCCGTATCGTCATGTTCACTGCT
 CTCCAACTGCATAGACAGTGTGTGTTGCAAGCAAGCTGATGAAGAATTCCAGATCCTGGC
 AAACCTCCTGGCGATACCTCCAGTGCATTACCAACAGGATATTTTTGCCATGGTGGATTTTG
 ATGAAGGCTCTGATGTATTTCAGATGCTAAACATGAATTCAAGCTCCAACCTTTCATCAACTTT
 CCTGCAAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGGTGCGGGGTTTTTCAGC
 TGAGCAGATTGCCCGGTGGATGCCGACAGAAGCTGATGTCATATTAGAGTGATTAGACCCC
 CAAATTATGCTGGTCCCTTATGTGAGGATTGCTTTTGGCTGTTATTGGTGGACTTGTGTAT
 CTTTGAAGAAGTAATATGGAATTTCTCTTAAATAAACTGGATGGGCTTTTGCAGCTTTGTG
 TTTTGTGCTTGTCTATGACATCTGGTCAAAATGTGGAACCATATAAGAGGACCACCATATGCCC
 ATAAGAATCCCCACACGGGACATGTAATTATATCCATGGAAGCAGTCAAGCCCAGTTTGTAT
 GCTGAAACACACATTGTTCTTCTGTTTAAATGGTGGAGTTACCTTAGGAATGGTGCCTTTTATG
 TGAAGCTGCTACCTCTGACATGGATATTGGAAGCGAAAGATAATGTGTGTGGCTGGTATTG
 GACTTGTGTATTATTTCTCAGTTGGATGCTCTCTATTTTTAGATCTAAATATCATGGCTAC
 CCATACAGCTTCTGTATGAGTTAAAGGTCCCAGAGATATATAGACACTGGAGTACTGGAA
 ATTTGAAAAACGAAATTCGTGTGTGTTTGAAGAAGAATGCAACTTGTATATTTTGTATTAC
 CTCCTTTTTTCAAGTGATTTAAATAGTTAATCATTTAACCAGAAAGATGTGTAGTGCCTTA
 ACAAGCAATCCTCTGTCAAAATCTGAGGTATTTGAAAATAAATTATCCTCTTAACTTCTCTT
 CCCAGTGAACCTTATGGAACATTTAATTTAGTACAATTAAGTATATTTAAAAAATTGTAAAA
 CTACTACTTTGTTTTAGTTAGAACAAGCTCAAACTACTTTAGTTAACTTGGTCATCTGAT
 TTTTATATTGCCTTATCCAAAGATGGGGAAGTAAGTCCTGACCAGGTGTTCCACATATGCC
 TGTTCAGATAACTATAGGAATTCATTCTTAGCTTCTTCATCTTTGTGTGGATGTGTAT
 ACTTTACGCATCTTTCCTTTTGTAGTAGAGAAATTATGTGTGTCATGTGGTCTCTGAAAAATG
 GAACACCAATCTTCAGAGCACAGCTCTAGCCCTCAGCAAGACAGTTGTTTCTCCTCCTCCTT
 GCATATTTCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCTCAAAAAAAGTA
 TCTCTAAATACAGGATTATAATTTCTGCTTGAGTATGGTGTTAACTACCTTGTATTTAGAAA
 GATTTAGATTCAATCCATCTCCTTAGTTTTCTTTAAGGTGACCCATCTGTGATAAAAAATA
 TAGCTTAGTGCTAAAAATCAGTGTAACCTATACATGGCCTAAAAATGTTTCTACAAATTAGAGT
 TTGTCATTATTCATTGTACCTAAGAGAAAAATAGGCTCAGTTAGAAAAAGGACTCCCTGG
 CCAGGCGCAGTGACTACGCCTGTAATCTCAGCACTTTGGGAGGCCAAGGCAGGCAGATCAC
 GAGGTGAGGAGTTCGAGACCATCTGGCCAACATGGTGAACCCCGTCTCTACTAAAAATAT
 AAAAAATTAGCTGGGTGTGGTGGCAGGAGCCGTGAATCCGAGCTACACAGGAGGCTGAGGCAC
 GAGAATCACTTGAACTCAGGAGATGGAGTTTCACTGAGCCGAGATCACGCCACTGCATCC
 AGCCTGGCAACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAA

FIGURE 74

MAARWRFWCVSVMVALLIVCDVPSASAQRKKEMVLSEKVSQLMEWTKRNPVIRMNGDKFR
 RLVKAPPRNYSVIVMFTALQLHRQCVCVKQADEEFQILANSWRYSSAFTNRIFFAMVDFDEG
 SDVFQMLNMNSAPTFINFPKGGKPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNY
 AGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFALCFVLAMTSGQMWNHIRGPPYAHKN
 PHTGHVNYIHGSSQAQFVAETHIVLLFNGGVTLMVLLCEAATSDMDIGKKRIMCVAGIGLV
 VLFFSWMLSIFRSKYHGYPYSFLMS

Signal peptide:

amino acids 1-29

Transmembrane domains:

amino acids 183-205, 217-237, 217-287, 301-321

FIGURE 75

FIGURE 76

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64885

<subunit 1 of 1, 536 aa, 1 stop

<MW: 61450, pI: 9.17, NX(S/T): 7

MLLLWVSVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDFSFGRLTFHFGSQVVKLPFFINF
MKTRGTSFLNAYTNSPICCPsRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGVRTQ
KFGKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPVNLIRNRTKVRVMERDWQNTDKAVNW
LRKEAINYTEPFVIYLG LNLPHYPSPSSGENFGSSTFHTSLYWLEKVS HDAIKIPKWSPLS
EMHPVDYYSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQLDLLQKTIVIIY
SSDHGELAMEHRQFYKMSMYEASAHVPLLMGPGIKAGLQVSNVSLVDIYPTMLDIAGIPL
PQNLSGYSLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYS DGAS
ILPQLFDLSSDPDEL TNVAVKFP EITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIG
QNYSNVIANLRWHQDWQKEPRKYENAI DQWLKTHMNPRAV

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.

amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,
498-501

Sulfatases proteins:

amino acids 286-315, 359-369, 78-97

FIGURE 77

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAG
 AGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**
 GCCTCTCTTGGCCTCCAACCTGTGGGCTACATCCTAGGCCCTTCTGGGGCTTTTGGGCACACT
 GGTGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAG
 CAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCAGGCATCACCAG
 TGTGACATCTATAGCACCTTCTGGGCTGCCCGTGACATCCAGGCTGCCAGGCCATGAT
 GGTGACATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCA
 CAGTCTTCTGCCAGGAATCCCGAGCCAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTC
 ATCCTTGGAGGCCTCCTGGGATTCATTCTCTGTGCCTGGAATCTTCATGGGATCCTACGGGA
 CTCTACTACCACTGGTGCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGG
 GCATTATTTCTTCCCTGTTCTCCTCGATAGCTGGAATCATCCTCTGCTTTTCTGCTCATCC
 CAGAGAAATCGCTCCAACCTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTC
 TCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCTACAGCCTGACAGGGT
 ATGTG**TGA**GAACCAAGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAG
 CACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGTGTCAGAAGGTGCTGCTGAGG
 ATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATG
 CAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTC
 CCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGG
 ATCCCTTTGCCCTCTGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCCTG
 ACTGACCCTCTGTGATCAAAGACCCCTCTCTCTGGCTGAGGTGGCTCTTAGCTCATTGCTGG
 GGATGGGAAGGAGAAGCAGTGCGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCC
 TCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTATGACTCCACAGTGTCC
 AGACTAATTTGTGCATGAACTGAAATAAAACCATCTACGGTATCCAGGGAACAGAAAGCAG
 GATGCAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 78

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64886
><subunit 1 of 1, 230 aa, 1 stop
><MW: 24549, pI: 8.56, NX(S/T): 1
MASLGLQLVGYYILGLLGLLGLTLVAMLLPSWKTSYVVGASIVTAVGFSKGLWMECATHSTGIT
QCDIYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVF
FILGGLLGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCS
SQRRNSNYDAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV
```

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

AGATTATCAGGAAATAAATAAGTGGTTTTTCCAATGTACACCTGTAAAA

FIGURE 80

MVPRIFAPAYVSVCLLLCPREVIAPAGSEPWLCQFAPRCGDKIYNPLEQCCYNDAIVSLSE
TRQCGPPCTFWPCFELCCLDSFGLTNDFFVVKLVQGVNSQCHSSPISSKCESRRRFP

Signal peptide:

amino acids 1-25

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
220

FIGURE 81

CTCCACTGCAACCACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATT
TTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGCCCCCATGACTCCTTACCT
GATGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCTGCAGCACTGTT
GCTATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAACTGCACCTTCAGA
GTCTGCTTTGAGCAGTGTGCCCTGGACCTTCATGGTGAAGCTGATAAACCAGAACTGCGA
CTCAGCCCGGACCTCGGATGACAGGCTTTGTCGCAGTGTGAGTGAATGGAACATCAGGGGAA
CGATGACTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTACCTGAGA
TCTGGGATGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCACTTCATT
CTGTGACCTGTCTGAGGCCACCTTGCAGTGCCTTGAGGAGGCCACAGGTCCCCTTCTAG
AATTCTGGACAGCATGAGATGCGTGTGCTGATGGGGGCCAGGGACTCTGAACCTCCTGAT
GACCCCTATGGCCAACATCAACCCGGCACACCCCAAGGCTGGCTGGGGAACCTTCACCCCT
TCTGTGAGATTTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAA
TAAATTTATGTACTTTATAAATGAAAA

FIGURE 82

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTFYLMLCQPHKRCGDKFYDPLQHCCYDDAVVP
LARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCD SARTSDDR LCRSVS

Signal peptide:

amino acids 1-24

Downloaded from www.ascp.com

FIGURE 83

GGGGGCGGGTGCCGTGGAGCACGGCGCTGGGGCCGCCCGCAGCGCTCACTCGCTCGCACTCAG
 TCGCGGGAGGCTTCCCGCGCCGGCGCGTCCCGCCCGCTCCCGGCGACCAGAAGTTCTCTT
 GCGCGTCCGACGGCGAC**ATGGG**CGTCCCCACGGCCCTGGAGGCCGGCAGCTGGCGCTGGGGA
 TCCCTGCTCTTCGTCTCTTCTGGCTGCGTCCCTAGGTCCGGTGGCAGCCTTCAAGGTGCG
 CACGCCGTATTCCCTGTATGTCTGTCCGAGGGGCGAAGCTCACCTCACCTGCAGGCTCT
 TGGGCCCTGTGGACAAAGGGCAGCATGTGACCTTCTACAAGACGTGGTACCGCAGCTCGAGG
 GCGGAGGTGCAGACCTGCTCAGAGCGCCGGCCCATCCGCAACCTCACGTTCCAGGACCTTCA
 CCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTGGCTCAGCGCCACGGGC
 TGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATGCGCAACCTGACCCTGCTG
 GATAGCGGCCTCACTGCTGCCTGGTGGTGGAGATCAGGCACCACCCTCGGAGCACAGGGT
 CCATGGTGCCATGGAGCTGCAGGTGCAGACAGGCAAGATGCACCATCCAACCTGTGTGGTGT
 ACCCATCCTCCTCCAGGATAGTGA**AA**ACATCAGGGCTGCAGCCCTGGCTACGGGTGCCTGC
 ATCGTAGGAATCCTCTGCCTCCCCCTCATCTGCTCCTGGTCTACAAGCAAAGGCAGGCAGC
 CTCCAACCGCGTGGCCAGGAGCTGGTGGGATGGACAGCAACATTCAAGGGATTGA**AA**ACC
 CCGGCTTTGAAGCCTCACCACTGCCAGGGGATACCCGAGGCCAAAGTCAGGCACCCCTG
 TCCTATGTGGCCAGCGGCAGCCTTCTGAGTCTGGCGGCATCTGCTTTCGGAGCCCCAGCAC
 CCCCCTGTCTCCTCCAGGCCCGGAGACGTCTTCTTCCCATCCTTGGACCTGTCCCTGACT
 CTC**AA**AACTTTGAGGTCATCT**AG**CCAGCTGGGGGACAGTGGGCTGTTGTGGCTGGGTCTGG
 GGCAGGTGCATTGTAGCCAGGGCTGGCTCTGTGAGTGGCCTCCTTGGCCTCGGCCCTGGTTC
 CCTCCCTCCTGCTCTGGGCTCAGATACTGTGACATCCAGAAGCCAGCCCTCAACCCCTC
 TGGATGCTACATGGGGATGCTGGACGGCTCAGCCCTGTTCCAAGGATTTTGGGGTGTGAG
 ATTCTCCCTTAGAGACCTGAAATTACCAGCTACAGATGCCAAATGACTTACATCTTAAGAA
 GTCCTAGAACGTCACAGCCCTTCAGCAGCTCTCGTTCTGAGACATGAGCCTTGGGATGTGGCA
 GCATCAGTGGGACAAGATGGACACTGGGCCACCCTCCAGGCACCAGACACAGGGCACGGTG
 GAGAGACTTCTCCCCGTGGCCGCCTTGGCTCCCCCGTTTTGCCCGAGGCTGCTCTTCTGTC
 AGACTTCTCTTTGTACCACAGTGGCTCTGGGGCCAGGCCTGCCTGCCACTGGCCATCGCC
 ACCTTCCCCAGCTGCCTCTACAGCAGTTTCTCTGAAGATCTGTCAACAGGTTAAGTCAAT
 CTGGGGCTTCCACTGCCTGCATCCAGTCCCAGAGCTTGGTGGTCCCGAAACGGGAAGTAC
 ATATTGGGGCATGGTGGCCTCCGTGAGCAAATGGTGTCTTGGGCAATCTGAGGCCAGGACAG
 ATGTTGCCCCACCCTGGAGATGGTGTCTGAGGGAGGTGGGTGGGGCCTTCTGGGAAGGTGA
 GTGGAGAGGGGCACCTGCCCCCGCCCTCCCATCCCCTACTCCACTGCTCAGCGCGGGCC
 ATTGCAAGGGTGCCACACAATGTCTTGTCACCCCTGGGACACTTCTGAGTATGAAGCGGGAT
 GCTATT**AAAA**ACTACATGGGG**AAAA**AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURE 84

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64897

><subunit 1 of 1, 311 aa, 1 stop

><MW: 33908, pI: 6.87, NX(S/T): 6

MGVPTALEAGSWRWGSLFLFALAASLGPVAAFKVATPYSLYVCPEGQNVTLTCRLLGPVDK
GHDVIFYKTWYRSSRGEVQTCSERRPIRNLTQQDLHLHHGGHQAANTSHDLAQRHGLASD
HHGNFSITMRNLTLLDSGLYCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNCCVYPSSSQ
DSENITAAALATGACIVGILCLPLILLIVYKQRQAASNRRAQELVRMDSNIQGIENPGFEAS
PPAQGIPEAKVRHPLSYVAQRQPSSESGRHLLSEPSTPLSPPGPGDVFFPSLDPVPDSPNFEVI

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 190-216

FIGURE 85

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCT
 TTCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCC
 TTTCTGCCCACCGTGCTTCTTGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGG
 GGTCTGTGGGTTGATCTGTGGCCCTGTGCCCTCCGTGTCTTTTCGTCTCCCTTCCTCCCGA
 CTCGCTCCCGACCAGCGGCTGACCTGGGAAAGGATGGTTCCCGAGGTGAGGTCCTC
 TCCTCCTTGCTGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCAGCTCGAGCCCGCC
 AGACATGTTCTGCCTTTTCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCTACT
 TGGAGCCACAAGGCCTGATGTA CTGCTGCGCTGTACCTGCTCAGAGGGCGCCATGTGAGT
 TGTACCGCCTCCACTGTCCGCTGTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATG
 CTGTCCCAAGTGTGTGAACCTCACACTCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCC
 AGCACAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCTCC
 CGCTGCCCAACCAAGTGTGTCCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCCTAC
 AACCTGCCCCGAACCAGGCTGCCCAGCACCCCTCCCACTGCCAGACTCCTGCTGCCAAGCCT
 GCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTG
 AGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGCCCGGGCACCCAGC
 CCCCCTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCACTTCAGACCCAAGGGAGCAG
 GCAGCACAACGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGG
 AAGACGTACTCCCACGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCCTTGCCCTG
 CATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGT
 ACCCTTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTGCCAGAGGACAAA
 GCAGACCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGTCTT
 CGTCCACACATCGGTATCCCAAGCCAGACAACCTGCGTCGCTTTGCCCTGGAACACAGAG
 CCTCGGACTTGGTGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAG
 AGAGGTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGA
 AAGTCAGGAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCAC
 GAAGGTCAGTGGAACGTCTTCTAGCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGA
 CAAAGTGACCAAGACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTT
 ATTATATATTAATAAATAAGAAGTTGCATTACCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 86

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902

><subunit 1 of 1, 451 aa, 1 stop

><MW: 49675, pI: 7.15, NX(S/T): 1

MVPEVRVLSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHYPYLEPQGLMYCLRCT
CSEGAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVPHTPSGLRAPPKSCQHNGTMYQHGEIF
SAHELFP SRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEED
SVQSLHGVRHPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKH
KKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRTCTEYPCRHPKVAGKC
CKICPEDKADPGHSEISSTRCPKAPGRVIVHTSVSPSPDNLRRFALEHEASDLVEIYLWKL
KDEETEAQRGEVFGPRPHSQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSDPG
AEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 87

AAAAAAAAAAAA

MDSLRKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLA
TLQEAATTQENVAWRKNWVMVGEGGASGRSP

Signal peptide:

amino acids 1-18

THE UNIVERSITY OF CHICAGO PRESS
50 EAST LAKE STREET, CHICAGO, ILL. 60607
TEL: 773-707-5000 FAX: 773-707-0868
WWW.CHICAGO.PRESS.EDU

[illegible]

FIGURE 90

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64905

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46038, pI: 6.50, NX(S/T): 2

MGPSTPLLLILFLLSWGPLQGQQHHLVEYMERRLAALAEERLAQCQDQSSRHAAELRDFKNKM
LPILLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGR
RNEKYDMVTDCGYTISQVRSMKILKRFGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRL
RDFTLAMAARKASRVRVFPWPVGTGQLVYGGFLYFARRPPGRPGGGGEMENTLQLIKFHLAN
RTVVDSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCCLAKLDPQTLDTQ
QWDTPCPRENAEAAFVICGTLVYVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAH
ASLRYNPRERQLYAWDDGYQIVYKLEMRKKEEV

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 177-180, 248-251

FIGURE 91

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAG
 AGCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTTCTCTCTCTAATCCAT
 CCGTCACTCTCTGTCTATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCC**ATGG**
 CTCTCATGCTCAGTTTGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTT
 GGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGAGGACGCAGCATTCTCTGTCTTCTGTCTC
 TCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTTCTCTAGCGTGG
 TCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGG
 ACAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTAC
 TGTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCACAGTCTTACTACCAGAAGGCCA
 TCTGGGAGCTACAGGTGTCTAGCACTGGGCTCAGTTTCTCTCATTTCATCACGGGATATGTT
 GATAGAGACATCCAGTACTCTGTCTAGTCCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTG
 GAAAGGTCACAAGGACAGGATTGTGCCACAGACTCCAGGACAAACAGAGACATGCATGGCG
 TGTGTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAACGCCGGGAGCATATCCTGTTCCATG
 CGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTTCGA
 GCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTATTTTTTG
 GCATTGTTGGACTGAAGATTTTCTCTCAAATTCCAGTGAAAAATCCAGGCGGAAGTGGAC
 TGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGAAACACGCAGTGGAGGTGAC
 TCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATA
 GAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCT
 TCTCAGAGTTTCCAAGCAGGGAACATTACTGGGAGGTGGACGGAGGACAAATAAAGGTG
 GCGCGTGGGAGTGTCCGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCG
 ATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTCACATTAAATCCCCGT
 TTTATCAGCGTCTTCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTATGAGTG
 TGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCTGACATGTCGGT
 TTGAAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAAGTCCC
 ATAGTCATCTGCCAGTCAACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGC
 AATCCAGAGACAAGCAACAGTGAGTCTCTCTCACAGGCAACACGCCTTCTCTCCCAGG
 GTGAAATG**TAG**GATGAATCACATCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCCA
 GATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCAGATGAAGGGGACTGGCCTGTCC
 ACATGGGAGTCAGGTGTCTAGGCTGCCCTGAGCTGGGAGGGAAGAGGCTGACATTACATTT
 AGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAATACCACTCTCAGGTGAAGAACCG
 TCAGGAATTTCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGC
 TTAGATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTA
 AAAAAA

FIGURE 92

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCLSPKTNAEAMEVRFFRGQFSS
 VVHLYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQK
 AIWELQVSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGQDLSTDSTRNDRMH
 GLFDVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLF
 FGIVGLKIFFSKFQWKIQAELEDWRRKHGQAELRDARKHAVEVTLDPETAHFKLCVSDLKTVT
 HRKAPQEVPHSEKRFRTRKSVVASQSFSQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLS
 PDHGYWVLRNLNGEHLVFTLNPRFISVFPTPTTKIGVFLDYECGTISFFNINDQSLIYTLTC
 RFEGLLRPYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAIPESTSNESSQATTPFLP
 RGEN

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 93

GCGATGGTGC GCCCGGTGGCGGTGGCGGCGCGGTTGCGGAGGCTTCTTGGTCGGATTGCA
 ACCAGGAGAAAGATGACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGC
 CATGAGGAGCCTGCCGAGCCTGGGCGGCCCTCGCCCTGTTGTGCTGCGCGCCGCCGCCGCCGCG
 CCGTCGCGCTCAGCCGCTCGGCGGGGAATGTACCAGGTGGCGGCGGGGCCGCGGGCAGGTG
 GACGCGTCGCCGGGCCCCGGGTTGCGGGGCGAGCCAGCCACCCCTTCCCTAGGGCGACGGC
 TCCACGGCCCCAGGCCGAGGACCGGGCCCCCGCGCGCCACCGTCCACCGACCCCTGGCTG
 CGACTTCTCCAGCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCCTCTTCC
 ACCACCTTTCAGGCGCGCTCGGCCCTCGCCGACCACCCCTCCGGCGGCGGAACGCACTTC
 GACCACCTCTCAGGCGCCGACCAGACCCGCGCCGACCACCCCTTCGACGACCACTGGCCCGG
 CGCCGACCACCCCTGTAGCGACCACCGTACC GGCGCCACGACTCCCCGGACCCCGACCCCG
 GATCTCCCCAGCAGCAACAGCAGCGTCTCTCCACCCCACTGCCACCGAGGCCCCCTC
 TTCGCTCTCCAGAGTATGTATGTAAGTCTGTGTTGGAAGCCTGAATGTGAATCGCT
 GCAACCAGACCAGGGCAGTGTGAGTGTGCGCCAGGTTATCAGGGGCTTCACTGTGAAACC
 TGCAAGAGGGCTTTTACCTAAATTACACTTCTGGGCTCTGTGAGCCATGTGACTGTAGTCC
 ACATGGAGCTCTCAGCATACCGTGCAACAGGTAAGCAACAGAGGTTGGAACCTGAAGTTTATT
 TTATTTTAGCAAGGGAAAAAAGGCTGCTACTCTCAAGGACCATACTGGTTTAAACAAAG
 GAGGATGAGGGTCATAGATTTACAAAATATTTATATACTTTTATCTCTTACTTTATATGT
 TATATTTAATGT CAGGATTTAAAAACATCTAATTTACTGATTTAGTTCCTCAAGCACTAG
 AGTCGCCAATTTTTCTGCGGATAATTTCTGTAAATTT CATGGGAAAAAATTATTGAAGAA
 AAATCTGCTTTCTGGAAGGGCTTTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTT
 ATGTTTATTAATATACCATTTGGAGTTT GAGGAAATTTGTTGTTTGGTTTATTTTTCTCTCTA
 ATCAAAATCTACATTTGTTTCTTTGGACATCTAAAGCTTAACTGGGGGTACCCTAATTTA
 TTTAACTAGTGGTAAGTAGACTGGTTTTACTCTATTTACCACTACATTTTGGAGCCAAAAG
 TAGATTAAGCAGGAATTATCTTTAAACTATTATGTTATTTGGAGGTAATTTAATCTAGTGGA
 ATAATGTACTGTTATCTAAGCATTGCGCTTGACTGCACTGAAAGTAATTTATCTTTGACCT
 TATGTGAGGCACCTTGGCTTTTTGTGGACCCCAAGTCAAAAACCTGAAGAGACAGTATTAAT
 AATGAAAAAATAATGACAGTTTATACTCAGTGTAACTGGGTATAACCCAAGATCTGCTGC
 CACTTACGAGCTGTGTTCTTGGCAAGTAATTTCTTTCACTGAGCTTGTTCCTCTCAAG
 GTTGTGTGAAGATTAATAGAGTTGATATATATAAAATGCCTAGCACATGTCACTCAATAAA
 TTCTGTTTGTTTAATTTCAAAGGAATATTATGGACTGAAATGAGAGAACATGTTTAAAGA
 ACTTTAGATCTCTTGACAAAAGATGCTTTATACTTTAGCACTAAATATTTTAAATGCTTTA
 TAAATGTATATACTGTTATGGAATATTGTATCATATTGTAGTTTATTAATAAGTAGAAG
 AGGCTGGGCGCGGTGGCTCAGCGCTGTAATCCTAGCACTTTGGGAGGCCAAGGCGGGTGGAT
 CACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAACCCCGTCTCTACTAAA
 AATACAAAACAAATGCTGGGCGTGGTGGCACACACCTGTAGTCCAGCTACTCGGGAGGCT
 GAGGCAGGAGAATCGGTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCTGAGATCGGCCACT
 GCACTCCAGCCTGGTGAGAGAGGGAGACTCTGTCTTAAAAAAAAAAAAAAAAAAAAAAAAA

10015671.121101

FIGURE 94

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952

><subunit 1 of 1, 258 aa, 1 stop

><MW: 25716, pI: 8.13, NX(S/T): 5

MRSLSLGLGGLALLCCAAAAAASASAASAGNVTGGGGAAGQVDASPGPGLRGEPSPHPFRATA
PTAQAPRTGPPRATVHRPLAATSPAQSPETTPWATAGPSSTTFQAPLGPSPPTPPAAERTS
TTSQAPTRPAPTTLSTTTTGAPTTVPATTVPAPTTPTPTDLPSSSNSSVLPTPPATEAPS
SPPPEYVCNCSVVGSLNVNRCNQTTGQCECRPGYQGLHCETCKEGFYLYNYSGLCQPCDCSP
HGALSIPCNR

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

EGF-like domain cysteine pattern signature.

amino acids 214-226.

FIGURE 95

TCGGGCGCAGTGTAGACCTGGGAGCA**ATG**GGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGG
 TCTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGG
 CCCTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAA
 GAACGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCCTCTC
 AGCACGGGCTGGGAGGGTGTGACCAGAGTGTATGGACCTGATAAAGCGAACTCCGGATGG
 GTGTTTGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGA
 CTATGCCATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGT
 ACAGTCTGACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTACCAAGTGGAGCAGGAGC
 CTGGGCTTCCTGTCACAG**TAG**CAGGCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAAGAT
 CCTTCTGTGAGTGTGCGTCCCCAGTAGGGATGGCGCCACAGGGTCCTGTGACCTCGGCCA
 GTGTCCACCCACCTCGCTCAGCGGCTCCCGGGGCCAGCACCAGCTCAGAATAAAGCGATTTC
 CACAGCA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
21

FIGURE 96

MGGLLLAFLALVSVPRQAQAVWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVT
LTPENNLRTLSSQHGLGGCDQSVMDLIKRNSGWVFENPSIGVLELWVLATNFRDYAIIFTQL
EFGDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

Signal peptide:

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 97

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGAC**ATG**CTGCTGCTGCTGCTGCC
 CCTGCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTT
 CCGTGACGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCAT
 GGTGATTTACCCCTGGCCAGTAGTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGA
 CCAGGATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACC
 GATTCACCTCCTTGGGGACCCACATACCAAGAATTGCACCTGAGCATCAGAGATGCCAGA
 AGAAGTGATGCGGGGAGATACTTCTTTCTGTATGGAGAAAGGAAGTATAAAATGGAATTATAA
 ACATCACCGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAG
 GCACCTGGAGTCCGGCTGCCCCAGAACTCTGACCTGCTCTGTGCCCTGGGGCTGTGAGCAG
 GGGACACCCCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCAC
 CCGCTCCTCGGTGCTCACCTCATCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTC
 AGGTGACCTTCCCTGGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCTTAC
 CCGCTCAGAACTTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTTGGG
 AAATGGCTCATCTCTGTCACTCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATG
 CAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCTGACCCTGTGCCCC
 TCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGA
 ATTACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGC
 AGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTGGGGGAGCTGGAGCCACAGCCCTG
 GTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAATCGGCAAG
 GCCAGCAGCGGGCTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTACGCCT
 CTCAGGGGCCCTGACTGAACCTTGGGCAGAAAGACAGTCCCCAGACCAGCCTCCCCAGCT
 TCTGCCCCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCGCTCAGCTTCCAGATGGT
 GAAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCC
 ACAGAT**TGAG**AAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGA
 GAAGTCAGAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACT
 ATGAATTATGTGCAGAGTGAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAAT
 CCACACTGTGCCCTCCCTTTTATTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 98

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWF
 REGANTDQDAPVATNNPARAVWEETRDRFHLGDPHTKNCTLSIRDARRSDAGRYFFRMEKG
 SIKWNYKHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTTPPMISWIGTSVS
 PLDPSTTRSSVLTLPQPQDHGTS LTCQVTFPGASVTNKT VHLNVSYPPQNLMTVFQGDG
 TVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWV
 HLRDAAEFTCRAQNPLGSQQVYLVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRS
 CRKKSARPAAGVGDGTGIEDANAVRGSSASQGPLEPWAEDSPFDQPPPASARSSVGEQELQYA
 SLFSQMVKFPWDSRGQEATDTEYSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 99

GACGCCCAGTGACCTGCCGAGGTCGGCAGCACAGAGCTCTGGAGATGAAGACCCTGTTCTCTG
GGTGTACGCTCGGCCTGGCCGCTGCCCTGTCTTCACCCCTGGAGGAGGAGGATATCACAGG
GACCTGGTACGTGAAGGCCATGGTGGTCGATAAGGACTTTCGGGAGGACAGGAGGCCAGGA
AGGTGTCCCCAGTGAAGGTGACAGCCCTGGGCGGTGGGAAGTTGGAAGCCACGTTACCTTC
ATGAGGGAGGATCGGTGCATCCAGAAGAAAATCCTGATGCGGAAGACGGAGGAGCCTGGCAA
ATACAGCGCCTATGGGGGAGGAAGCTCATGTACCTGCAGGAGCTGCCAGGAGGGACCACT
ACATCTTTTACTGCAAAGACCAGACCATGGGGGCTGCTCCACATGGGAAAGCTTGTGGGT
AGGAATTCTGATACCAACCGGGAGGCCCTGGAAGAATTTAAGAAATTGGTGCAGCGCAAGGG
ACTCTCGGAGGAGGACATTTTCACGCCCTGCAGACGGGAAGCTGCGTTCCCGAACACTAGG
CAGCCCCCGGGTCTGCACCTCCAGAGCCACCCCTACCACCAGACACAGAGCCCGGACCACCT
GGACCTACCCTCCAGCCATGACCCTTCCTGTCTCCACCCACCTGACTCCAAATAAAGTCCT
TTTCCCCCAA

FIGURE 100

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65404

<subunit 1 of 1, 170 aa, 1 stop

<MW: 19457, pI: 9.10, NX(S/T): 0

MKTLFLGVTGLAAALSFTLEEDITGTWYVKAMVVDKDFPEDRRPRKVSPVKVTALGGGKL

EATFTFMREDRCIQKKILMRKTEEPGKYSAYGGRKLMYLQELPRRDHYIFYCKDQHHGGLLH

MGKLVGRNSDTNREALLEEFKKLVQRKGLSEEDIFTPLQTGSCVPEH

Important features:**Signal peptide:**

amino acids 1-17



FIGURE 101

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCAC
 AGCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCC**ATG**AGGATTCTGCAGTTAA
 TCCTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTTCGAG
 TGCAAGCCTCACTCCAGCCCTGGCAGGCAGCCCTGTTTCGAGAAGACGCGGCTACTCTGTGG
 GCGCAGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACA
 TAGTTCACCTGGGGCAGCACAACTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCC
 ACTGAGTCCTTCCCCACCCGGCTTCAACAACAGCCTCCCCAACAAGACCACCGCAATGA
 CATCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCC
 TCTCCTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCC
 AGCCCCAGTTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTTGAGCACCA
 GAAGTGTGAGAACGCCTACCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGG
 AAGGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCACTCTCTT
 CAAGGCATTATCTCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACAC
 GAAAGTCTGCAATATGTGGACTGGATCCAGGAGACGATGAAGAACAAT**TAG**ACTGGACCCA
 CCCACCACAGCCCATCACCTCCATTTCCACTTGGTGTGGTTCCTGTTCCTCTCTGTTAAT
 AAGAAACCTTAAGCCAAGACCCTCTACGAACATTCTTTGGGCCCTCCTGGACTACAGGAGATG
 CTGTCACTTAATAATCAACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTT
 GAAATATTGTGACTCTGGGAATGACAACACCTGGTTTGGTTCTCTGTGTATCCCCAGCCCCA
 AAGACAGTCCTGGCCATATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 102

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65405

<subunit 1 of 1, 250 aa, 1 stop

<MW: 27466, pI: 8.87, NX(S/T): 4

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWILLTAAH
 CLKPRYIVHLGQHNLKKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITW
 AVRPLTLSSRCVTAGTSCSLISGWGSTSSPQLRLPHTLRCANITIIIEHQKCNAYPGNITDTM
 VCASVQEGGKDSCQGDGSGGPLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIQETMKNN

Important features:**Signal peptide:**

amino acids 1-18

Serine proteases, trypsin family, histidine active site.

amino acids 58-63

N-glycosylation sites.

amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.

amino acids 145-148

Kringle domain proteins.

amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein

amino acids 199-209, 47-63, 220-243

Apple domain proteins

amino acids 222-249, 189-222

FIGURE 103

GAGCAGTGTCTGCTGGAGCCG**ATG**CCAAAAACCATGCATTCTTATTTCAGATTCAATTGTTT
TCTTTTATCTGTGGGGCCTTTTTACTGCTCAGAGACAAAAGAAAGAGGAGAGCACCGAAGAA
GTGAAAATAGAAGTTTTGCATCGTCCAGAAAACCTGCTCTAAGACAAGCAAGAAGGGAGACCT
ACTAAATGCCATTATGACGGCTACCTGGCTAAAGACGGCTCGAAATTCTACTGCAGCCGGA
CACAAAATGAAGGCCACCCCAAATGGTTTGTCTTGGTGTGGGCAAGTCATAAAAGGCTA
GACATTGCTATGACAGATATGTGCCCTGGAGAAAAGCGAAAAGTAGTTATACCCCTTCATT
TGCATACGGAAGGAAGGCTATGCAGAAGGCAAGATTCCACCGGATGCTACATTGATTTTTG
AGATTGAACTTTATGCTGTGACCAAAGGACCACGGAGCATTGAGACATTTAAACAAATAGAC
ATGGACAATGACAGGCAGCTCTCTAAAGCCGAGATAAACCTCTACTTGCAAAGGGAATTTGA
AAAAGATGAGAAGCCACGTGACAAGTCATATCAGGATGCAGTTTTAGAAGATATTTTAAGA
AGAATGACCATGATGGTGATGGCTTCATTTCCTCCCAAGGAATACAATGTATACCAACACGAT
GAACTA**TAG**CATATTTGTATTTCTACTTTTTTTTTTTTAGCTATTTACTGTACTTTATGTATA
AAACAAAGTCACTTTTCTCCAAGTTGTATTTGCTATTTTCCCTATGAGAAGATATTTTGA
TCTCCCCAATACATTGATTTTGGTATAATAAATGTGAGGCTGTTTTGCAAACCTTAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA


FIGURE 104


```


</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65406
<subunit 1 of 1, 222 aa, 1 stop
<MW: 25794, pI: 6.24, NX(S/T): 1
MPKTMHFLFRFIVFFYLVGLFTAQRQKKEESTEEVKIEVLHRPENCSKTSKKGDLLNAHYDG
YLAKDGSKFYCSRTQNEGHPKWFVLGVGQVIKGLDIAMTDCPGEKRKVVIPPSFAYGKEGY
AEGKIPPDATLIFEIELYAVTKGERSIETFKQIDMDNDRQLSKAEINLYLQREFEKDEKPRD
KSYQDAVLEDFKKNDDHGDGFISPKYENVYQHDEL


```


Important features:


 **Endoplasmic reticulum targeting sequence.**


 amino acids 219-222

 **N-glycosylation site.**

 amino acids 45-48

 **FKBP-type peptidyl-prolyl cis-trans isomerase**

 amino acids 87-223, 129-142

 **EF-hand calcium-binding domain proteins**

amino acids 202-214, 195-214

AGTACTTATATTCTC

FIGURE 106

MQGPIIIIPGLCFLLSLFGAVTQKTKTSCAKCPPNASCVNTHCTCNHGYTSGSGQKLF¹TFPL
ETCNARHGGSRL

Signal peptide:

amino acids 1-18

106/249

FIGURE 107

CAAGCAGGTCATCCCCTTGTTGACCTTCAAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGCAC
AGGAAAAGGGTGACCTCTGAGATTTCCCTTTTCCCCCAGACTTTGGAAGTGACCCACC**ATGG**
GGCTCAGCATCTTTTTGCTCCTGTGTGTTCTTGGGCTCAGCCAGGCAGCCACACCGAAGATT
TTCAATGGCACTGAGTGTGGGCGTAACCTCACAGCCGTGGCAGGTGGGGCTGTTTGAGGGCAC
CAGCCTGCGCTGCGGGGGTGCTTATTGACCACAGGTGGGTCTCTACAGCGGCTCACTGCA
GCGGCAGCAGGTAAGTGGGTGCGCCTGGGGGAACACAGCCTCAGCCAGCTCGACTGGACCGAG
CAGATCCGGCACAGCGGCTTCTCTGTGACCCATCCCGGTACCTGGGAGCCTCGACGAGCCA
CGAGCACGACCTCCGGCTGCTGCGGCTGCGCCTGCCCCTCCGCGTAACCAGCAGCGTTCAAC
CCCTGCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGCCACGCTCTCAGGCTGGGGC
ATCACCAACCACCCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGT
CTCCCATGCCACCTGCCATGGTGTGTATCCCGGGAGAATCACGAGCAACATGGTGTGTGCAG
GCGGCGTCCCGGGGCAGGATGCCCTGCCAGGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGGA
GTCCTTCAAGGTCTGGTGTCTTGGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCCCTGG
AGTCTACACCTATATTTGCAAGTATGTGGACTGGATCCGATGATCATGAGGAACA**CTGAC**
CTGTTTCCTCCACCTCCACCCCAACCCCTTAACCTTGGGTACCCCTCTGGCCCTCAGAGCACC
AATATCTCCTCCATCACTTCCCCTAGCTCCACTCTTGTGGCCTGGGAACCTCTTGGAACCTT
TAACTCCTGCCAGCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGA
ATAAATATAAATGAAGGAGGGGCAAAAAAAAAAAAAA

FIGURE 108

MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLTAAH
CSGSRYWVRLGEHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLRLPVRVTSSV
QPLPLPNDCATAGTECHVSGWGITNHPRNPFDDLQCLNLSIVSHATCHGVYPGRITSNMVC
AGGVPGQDACQGDSGGPLVCGGVLQGLVSWGSGVPCGQDGIPGVYTYICKYVDWIRMIMRNN

Signal peptide:

amino acids 1-17

108/249

FIGURE 109

GCGGCCACACGCAGCTAGCCGGAGCCCGGAGCCAGGCGCCTGTGCCTCCTCCTCGTCCCTCGC
 CGCGTCCGCGAAGCCTGGAGCCGGCGGGAGCCCCGCGCTCGCC**ATG**TCGGGCGAGCTCAGCA
 ACAGGTTCCAAGGAGGGAAGGCGTTTCGGCTTGCTCAAAGCCCCGCGAGGAGAGGAGGCTGGCC
 GAGATCAACCGGAGTTTCTGTGTGACCAGAAGTACAGTGATGAAGAGAACCTTCCAGAAAA
 GCTCACAGCCTTCAAAGAGAAGTACATGGAGTTTGACCTGAACAATGAAGGCGAGATTGACC
 TGATGTCTTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCCAAGACCCACCTGGAGATGAAG
 AAGATGATCTCAGAGGTGACAGGAGGGGTCACTGACACTATATCCTACCGAGACTTTGTGAA
 CATGATGCTGGGGAACGGTCGGCTGTCTCTCAAGTTAGTCATGATGTTTGAAGGAAAAGCCA
 ACGAGAGCAGCCCCAAGCCAGTTGGCCCCCTCCAGAGAGAGACATTGCTAGCCTGCCCT**TGA**
 GGACCCCGCTGGACTCCCCAGCCTTCCACCCCATACCTCCCTCCCGATCTTGCTGCCCTT
 CTTGACACACTGTGATCTCTCTCTCTCATTTGTTTGGTCATTGAGGGTTTGGTTTGTGTTT
 TCATCAATGTCTTTGTAAAGCACAAATTATCTGCCTTAAAGGGGCTCTGGGTGCGGGAATCC
 TGAGCCTTGGGTCCCCCTCCCTCTCTTCTTCCCTCCTTCCCCGCTCCCTGTGCAGAAGGGCTG
 ATATCAAACCAAACTAGAGGGGCGAGGGCCAGGGCAGGGAGGCTTCCAGCCTGTGTCTCC
 CTCACCTTGAGGAACACGACTCTCCATCCTTTCAGAAAGTCTCCAAGCCAAGTTCAGGCTC
 ACTGACCTGGCTCTGACGAGGACCCAGGCCACTCTGAGAAGACCTTGAGTAGGGACAAGG
 CTGCAGGGCCTCTTTCGGGTTTCCTTGACAGTGCCATGGTTCCAGTGCTCTGGTGTACCC
 AGGACACAGCCACTCGGGGCCCCGCTGCCCCAGCTGATCCCCACTCATCCACACCTTCTCT
 CATCCTCAGTGATGTGAAGTGGAAGGAAAGGAGCTTGGCATTTGGGAGCCCTTCAAGAAGG
 TACCAGAAGGAACCTCCAGTCCTGCTCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAG
 CGTGCAGCCCTACTGTCCCTTACTGGGGCAGCAGAGGGCTTCGGAGGCAGAAGTGAGGCCTG
 GGGTTTGGGGGGAAGGTCAGCTCAGTGCTGTTCACCTTTTAGGGAGGATACTGAGGGGAC
 CAGGATGGGAGAATGAGGAGTAAAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGA
 CTGAGAAATACAAGGTTGCTTGCTGTGACCCAATCTGCTTGAIAAAAAAAAAAAAAAAAAA

FIGURE 110

MSGELSNRFQGGKAFGLLKARQERRLAEGINREFLCDQKYSDEENLPEKLTAFKEKYMEFDLN
NEGEIDLMSLKRMMEKLGVPKTHLEMKKMISEVTGGVSDTISYRDFVNMMLGKRSAVLKLV
MFEGKANESSPKPVGPPPERDIASLP

005420

FIGURE 111A

CGCGCTCCCCGCGCGCCTCCTCGGGCTCCACGCGTCTTGCCCCGAGAGGCAGCGCTCCTCCA
GGAGCGGGGCCCTGCAACAC**ATG**CCCCCGGGTGGGCAGGGGTTCGGCGCGCCGTGCGCGCC
CGCCTGGCGCTGGCCTTGGCGCTGGCAGCGTCTCTGAGTGGGCCTCCAGCCGTGCCTGGCC
CACCAGTGTACCTGCTCCGCTGCCAGCTGGACTGCCACGGGCTCCGCGCGGCTTCCGCGGCTT
CTCGGGGATCCCCCGCAACGCTGAGCGCCTTGACCTGGACAGAAATAATATCACCAGGATC
ACCAAGATGGACTTCTGCTGGCTCAAGAACCTCCGAGTCTTGACACTTGGAGACAACAGGT
CAGCGTCATCGAGAGAGGCGCCTTCCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACA
AGAATAAGCTGCAAGTCTTCCAGAAATGCTTTCCAGAGCAGCGCGAAGCTCACCAGACTA
GATTTGAGTGAACACAGATCCAGGGGATCCCGAGGAAGGCGTTCCGCGGCATCACCAGTGT
GAAGAACCTGCAACTGGACAACAACCATCATCAGTGCATTGAAGATGGAGCCTTCCGAGCGC
TGCGCGATTTGGAGATCCTTACCCTCAACAACAACAACATCAGTCGCATCCTGGTCACCAGC
TTCAACCACATGCCGAAGATCCGAACCTCTGCGCCTCCACTCCAACCACCTCTACTGCGACTG
CCACTGGCCTGGCTCTCGGATTTGGCTGCGACAGCGACGGACAGTTGGCCAGTTTACACTCT
GCATGGCTCTCTGTCATTGAGGGGCTTCAACGTGGCGGATGTGCAGAAGAAGGAGTACGTG
TGCCAGCCCCCACTCGAGCCCCCATCTGCAATGCCAATCCATCTCTGCGCCTTCGCC
CTGCACGTGACAGCAATAACATCGTGGACTGTGAGGAAGGGCTTGATGGAGATTCCTGCCA
ACTTCCGGAGGGCATCTGTCGAAATACGCTAGAACAGAACTCCATCAAAGCCATCCCTGCA
GGAGCTCTCACCAGTACAGAAACTGAAGCGAATAGACATCAGCAAGATCAGATATCGGA
TATTGCTCAGATGCTCTTCAGGGCCTGAAATCACTCACATCGCTGGTCTGTATGGGAACA
AGATCACCAGAGATGCCAAGGGACTGTTTGATGGGCTGGTGTCCCTACAGCTGCTCCTCCTC
AATGCCAACAGATCACTGCTGCGGGTGAACACGTTTCAGGAGCTGACAGAACCTCACAATT
GCTCTCCCTGTATGACAACAGCTGACAGCCATCAGCAAGGCTCTTCGCCCTCTCGACT
CCATCCAGACACTCCACTTAGCCCAAACCCATTTGTGTGCGACTGCCACTTGAAGTGGCTG
GCCGACTACCTCCAGACAACCCCATCGAGACAAGCGGGGCCGCTGACGACGCCCGCGCCG
ACTCGCCAAACAGCGCATCAGCCAGATCAAGAGCAAGAAGTTCCGCTGCTCAGGCTCCGAGG
ATTACCGCAGCAGGTTTCCAGAGCAGTGTCTCATGGACCTCGTGTGCCCCGAGAAAGTGTGCG
TGTGAGGGCACGATTGTGGACTGCTCCAACCAGAAGCTGGTCCGCATCCCAAGCCACCTCCC
TGAATATGTCACCGAAGCTGGGACTGAATGACAATGAGGTATCTGTTCTGGAGGCCACTGGCA
TCTTCAAGAAGTTGCCAACCTGCGGAAAAATAAATCTGAGTAACAATAAGATCAAGGAGGTG
CGAGAGGAGCTTTGATGGAGCAGCCAGCGTGCAGGAGCTGATGCTGACAGGGAACCGATT
GGAGACCTGTCACGGGCGCGTGTTCGCTGGCCTCAGTGGCCTCAAACCTTGTATGCTGAGGA
GTAACCTGATCAGCTGTGTGAGTAATGACACCTTTGCGGCGCTGAGTTCCGTTGAGACTGCTG
TCCCTCTATGACAATCGGATCACCACCATCACCCCTGGGGCCTTACCACGCTTGTCTCCCT
GTCCACCTATAACCTCCTGTCCACCCCTTCAACTGCAACTGCCACTGGCCTGGCTCGGCA
AGTGTTTGAAGGAAGCCAGATCGTCAGTGGGAACCTAGTGCCAGAGAACCATTTTTCTC
AAGGAGATTTCCATCCAGGATGTGGCCATCCAGGACTTCACTGTGATGGCAACGAGGAGAG
TAGCTGCCAGCTGAGCCCGCGCTGCCGGAGCAGTGACCTGCTGAGAGACAGTGGTGCAT
GCAGCAACAAGGGGCTCCGCGCCTCCCCAGAGGCATGCCAAGGATGTGACCGAGCTGTAC
CTGGAAGGAAACCACTAACAGCCGTGCCAGAGAGCTGTCCGCCCTCCGACACCTGACGCT
TATTGACCTGAGCAACAACAGCATCAGCATGCTGACCAATTACACCTTCACTGATCTCTC
ACCTCTCCACTCTGATCTGAGCTACACCCGCTGAGGTGCATCCCGTCCACGCCCTTCAAC
GGGCTGCGGTCCCTGCGAGTGCTAACCTCCATGGCAATGACATTTCCAGCGTCTCTGAAGG
CTCCTTCAACGACCTCACACTCTTTTCCCATCTGGCGCTGGGAACCAACCACTCCACTGTG
ACTGCACTTCTCGGTGGCTTCTCGGAGTGGGTGAAGGCGGGGTACAAGGACCTGGCATCGCC
CGCTGCAAGTGGCTGAGCCCTGAGGCTGACAGGCTCCTGCTCACCACCAACCAACCGCTT
CCAGTGCAAGGGCCAGTGGACATCAACATTGTGGCCAAATGCAATGCCTGCCTCTCCAGCC
CGTGCAAGAATAACGGGACATGCACCCAGGACCTGTGGAGCTGTACCGCTGTGCTGCCCC

FIGURE 111B

TACAGCTACAAGGGCAAGGACTGCACTGTGCCCATCAACACCTGCATCCAGAACCCCTGTCA
 GCATGGAGGCACCTGCCACCTGAGTGACAGCCACAAGGATGGGTTACAGCTGCTCCTGCCCTC
 TGGGCTTTGAGGGGCAGCGGTGTGAGATCAACCCAGATGACTGTGAGGACAACGACTGCGAA
 AACAAATGCCACCTGCGTGGACGGGATCAACAACACTACGTGTGTATCTGTCCGCCTAACTACAC
 AGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTCTGTGACGATG
 AGGCCAAGTGCATCCCCCTGGACAAAGGATTACAGTGCAGAGTGTGTCCCTGGCTACAGCGGG
 AAGCTCTGTGAGACAGACAATGATGACTGTGTGGCCCAAGTGCCGCCACGGGGCCCAAGTG
 CGTGGACACAATCAATGGCTACACATGCACCTGCCCCAGGGCTTCAGTGGACCCCTTCTGTG
 AACACCCCCACCCATGGTCCTACTGCAGACCAGCCCATGCGACCACTACGAGTGCCAGAAC
 GGGGCCCAAGTGCATCGTGGTGCAGCAGGAGCCACCTGCCGCTGCCACCAGGCTTCGCGCG
 CCCCAGATGCGAGAAGCTCATCTGTCAACTTCGTGGGCAAGACTCCTACGTGGAAGTGG
 CCTCCGCCCAAGGTCGCCACCCAGGCCAACATCTCCCTGCAGGTGGCCACTGACAAGGACAAC
 GGCATCCTTCTCTACAAGGAGACAATGACCCCCCTGGCACTGGAGCTGTACAGGGGCCACGT
 GCGGCTGGTCTATGACAGCCTGAGTTCCCTCCAACCACAGTGTACAGTGTGGAGACAGTGA
 ATGATGGGCAGTTTACAGTGTGGAGCTGGTGACGCTAAACCAGACCTGAACCTAGTAGTG
 GACAAAGGAACCTCAAAGAGCCTGGGGAAGCTCCAGAAGCAGCAGCAGTGGGCATCAACAG
 CCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCTCTCCGCTTGGCCACAGGACGCG
 ACCGGCTCTAGGCGGCTTCCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAG
 GACTTCAAGGCCCTCCACCACAGTCCCTGGGGGTGTACCAAGGCTGCAAGTCTCTGCACCGT
 GTGCAAGCACGGCTGTGCCGCTCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGCCGCCACG
 GCTGGACCGGCCACTCTGCGACAGGAGGCCCGGGACCCCTGCCTCGGCCACAGATGCCAC
 CATGGAAAATGTGTGGCAACTGGGACCTCATACATGTGCAAGTGTGCCGAGGGCTATGGAGG
 GGACTTGTGTGACAACAAGAAATGACTTGCCAATGCCTGCTCAGCCTTCAAGTGTACCATG
 GGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCTGTGCCAGCCCGCTTTAGCGGC
 GAGCACTGCCAACAAGAGAATCCGTGCCTGGGACAAGTAGTCCGAGAGGTGATCCGCCGCCA
 GAAAGGTTATGCATCATGTGCCACAGCCTCCAAGGTGCCCATCATGGAATGTCGTGGGGCT
 GTGGGCCCAAGTGTGTCAGGCCACCCGCAGCAAGCGGCGAAATACGTCTTCCAGTGCACG
 GACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACCTTAGAGTGGCGTGCCTCGCGTGTTC
 CTAAGCCCCCTGCCGCTGCCTGCCACCTCTCGGACTCCAGCTTGATGGAGTTGGGACAGCC
 ATGTGGGACCCCTGGTGATTGACATGAAGGAAATGAAGCTGGAGAGGAAGGTAAAGAAGA
 AGAGAATATTAAAGTATATTGTAAATAAACAATAAATAAGAACTAAAAAAAAAAAAAAAAA
 AAAAAA

FIGURE 112

MAPGWAGVGAAVRARLALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRN
 AERLDLDNRNITRITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRNLNKNLQVL
 PELLFQSTPKLTRDLSENQIQGIPRKAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEIL
 TLNNNNISRIIVTSFNHMPKIRTLRLHSNHLVYCDCHLAWLSDWLRQRRTVGQFTLCMAPVHL
 RGFNVADVQKKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLMEIPANLPBGIV
 EIRLEQNSIKAI PAGAFQYKKLKRIDISKNIISDIAPDAFQGLKSLTSLVLYGNKITEIAK
 GLFDGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTIKGLFAPLQSIQTLHL
 AQNPFVCDCHLKWLDYLDQNPITSGARCSSPRRLANKRISQIKSKKFRCSGSEDYRSRFS
 SECFMDLVCPKEKCRCEGTIVDCSNQKLVRI PSHLPEYVTDRLRNDNEVSVLEATGIFKKLPN
 LRKINLSNNKI KEVREGAFDGAASVQELMLTGNQLETVHGRVFRGLSGLKTLMLRSNLISCV
 SNDTFAGLSSVRLLSLYDNRIITITPGAFTTLVSLSTINLLSNPFNCNCHLAWLGKWLKRRR
 IVSGNPRCQKPFFLKEIPIQDVAIQDFTCDGNEESSCQLSPRCFEQCTCMETVVRCSENKGLR
 ALPRGMPKDVTELYEGNHLTAVPRELSALRHLLTIDLNNISMLTNYTFNNMHLSTLIL
 SYNRLRCIPVHAFNGLRSLRVLTLLHGNDISSVPEGSFNDLTSLSHLALGNTNPLHCDCLRWL
 SEWVKAGYKEPGIARCSSPEPMADRLLLTTPTHRFQCKGPVDINIVAKCNACLSSPCKNNGT
 CTQDPVELYRCACPYSGKDKCTVPINTCIQNPCQHGGTCHLSDSHKDGFS CSCPLGFEGQR
 CEINPDCCEDNDCENNATCVDGINNYVICPPNYTGELCDEVIDHCVPENLNCQHEAKCIPL
 DKGFSCCEVPGYSGKLCETDNDDCAHKCRHGAQCVDTINGYCTCQGGFSGPFCEHPPPMV
 LLQTSPCDQYECQNGAQCI VVQQEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRP
 QANISLQVATDKDNGILLYKGDNDPLALELYQGHVRLVYDSLSPPTTVYSVETVNDGQFHS
 VELVTNLQTLNLVVDKGTPKSLGKLQKQPAVGINSPLYLGGIPTSTGLSALRQGTDRPLGGF
 HGCIHEVRINNELQDFKALPPQSLGVSPGCKSCTVCKHGLCRSVEKDSVVECECRPGWTGPLC
 DQEARDPCLGRCHHGKCVATGTSYMKCKAEGYGGDLCDNKNDSANACSAFKCHHGQCHISD
 QGEPYCLCQPGFSGEHCQENPCLGQVVREVI RRQKGYASCATASKVPIMECRGGCGPQCCQ
 PTRSKRRKYVFQCTDGSSEFVEEVERHLECGCLACS

Signal peptide:

amino acids 1-27

FIGURE 113

GGATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTCCTCCGCA
 GACTCAACTGAGAAGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCT
 CCGGCAGGCTTTGAGG**ATGA**AGGCTGCGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAG
 GCGCCGAGTCCAAAATCTACACTCGTTGCAAACCTGGCAAAAATATTCTCGAGGGCTGGCCTG
 GACAATTACTGGGGCTTCAGCCTTGAAACTGGATCTGCATGGCATATTATGAGAGCGGCTA
 CAACACCACAGCCCCGACGGTCCTGGATGACGGCAGCATCGACTATGGCATCTTCAGATCA
 ACAGCTTCGCGTGGTGACAGCGGAAAGCTGAAGGAGAACAACCACTGCCATGTCGCCTGC
 TCAGCCTTGATCACTGATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGA
 GACACAAGGAATGAACTATTGGCAAGGCTGGAAGAAACATTGTGAGGGCAGAGACCTGTCCG
 AGTGGA AAAAAGGCTGTGAGGTTTCCT**TAA**ACTGGAACCTGGACCCAGGATGCTTTGCAGCAAC
 GCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTATCTTGTCCCGTTTCCTCCCAATA
 TTCCTTCTCAAACCTGGAGAGGGAAAAATTAAGCTATACTTTTAAGAAAATAAATATTTCCAT
 TTAAATGTC

FIGURE 114

MKAAGILTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFS LGNWICMAYYESGYNTTAP
TVLDDGSIDYGIFQINSFAWCRRGKLKENNHCHVACSALITDDLTD AII CARKIVKETQGMN
YWQGWKKKHCEGRDLSEWKKGCEVS

Signal peptide:

amino acids 1-19

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 115

CAGGCCATTTGCATCCCACTGTCCTTGTGTTTCGGAGCCAGGCCACACCGTCCTCAGCAGTGT
 CATGTGTTAAAAACGCCAAGTCGAATATATCATGCCCCCTATTAATACTTGTACATGGCTCCC
 CATTGGTTTTTGGAGAAAAGTTCAAGCTTTTACCTTGGTGTCTGCCTGTATCCCAGTGTTC
 AGGCTGGCTAGACGGCGGAAGAAGATCCTATTTTACTGTCACTTCCCAGATCTGCTTCTCAC
 CAAGAGAGATTCTTTTCTTAAACGACTATACAGGGCCCCAATTGACTGGATAGAGGAATACA
 CCACAGGCATGGCAGACTGCATCTTAGTCAACAGCCAGTTCACAGCTGCTGTTTTTAAGGAA
 ACATTCAAGTCCCTGTCTCACATAGACCCTGATGTCTCTATCCATCTCTAAATGTCACCAG
 CTTTGACTCAGTTGTTCTCTGAAAAGCTGGATGACCTAGTCCCCAAGGGGAAAAAATTCCTGC
 TGCTCTCCATCAACAGATACGAAAGGAAGAAAAATCTGACTTTGGCACTGGAAGCCCTAGTA
 CAGCTGCGTGGAAGATTGACATCCCAAGATTGGGAGAGGGTTCATCTGATCGTGGCAGGTGG
 TTATGACGAGAGAGTCTGGAGAATGTGGAACATTATCAGGAATTGAAGAAAAATGGTCCAAC
 AGTCCGACCTTGGCCAGTATGTGACCTTCTTGAGGTCTTCTCAGACAAACAGAAAAATCTCC
 CTCTCCACAGCTGCACGTGTGTGCTTTACACACCAAGCAATGAGCACTTTGGCATTGTCCC
 TCTGGAAGCCATGTACATGCAGTGCCCAGTCATTGCTGTTAATTCGGGTGGACCCCTGGAGT
 CCATTGACCACAGTGTACAGGGTTTCTGTGTGAGCCTGACCCGGTGCACTTCTCAGAAGCA
 ATAGAAAAGTTTCATCCGTGAACCTTCTTAAAAGCCACCATGGGCCTGGCTGGAAGAGCCAG
 AGTGAAGGAAAAATTTTCCCCTGAAGCATTACAGAACAGCTCTACCGATATGTTACCAAAC
 TGCTGGTATTAATCAGATTGTTTTTAAGATCTCCATTAATGTCATTTTTATGGATTGTAGACC
 CAGTTTTGAAACCAAAAAAGAAACCTAGAATCTAATGCAGAAGAGATCTTTAAAAAATAAA
 CTTGAGTCTTGAATGTGAGCCACTTTCCTATATACCACACCTCCCTGTCCACTTTTCAGAAA
 AACCATGTCTTTTATGCTATAATCATTCCAAATTTTGCCAGTGTTAAGTTACAAATGTGGTG
 TCATTCCATGTTTCAGCAGAGTATTTTAATTATATTTTCTCGGGATTATTGCTCTTCTGTCTA
 TAAATTTTGAATGATACTGTGCCTTAATTGGTTTTTCATAGTTTAAAGTGTGTATCATTATCAA
 AGTTGATTAATTTGGCTTCATAGTATAATGAGAGCAGGGCTATTGTAGTTCACAGATTCAT
 CCACCGAAGTGTTCACTGTCTATCTGTTAGGGAATTTTGTGTTGCTGCTTTGCTGCGATC
 CATAGCGAGAGTCTCTGTATTTTTTTAAGATAATTTGTATTTTTGTCACACTGAGATATAA
 TAAAAGGTGTTTATCATAAAAA

FIGURE 116

MPLLKLVHGSPLVFGEKFKLFTLVSACIPVFRLARRRKILFYCHFPDLLLTKRDSFLKRLY
RAPIDWIEEYTTGMADCILVNSQFTA AVFKETFKSLSHIDPDVLYPSLNVTSFDSVVPEKLD
DLVPKGKKFLLLSINRYERKKNLTLALEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVE
HYQELKKMVQQSDLGQYVTF LRSFSDKQKISLLHSCTCVLYTPSNEHFGIVPLEAMYMQCPV
IAVNSGGPLESIDHSVTGFLCEPDVHFSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAF
TEQLYRYVTKLLV

Signal peptide:

amino acids 1-15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 117

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACCATGTTGGACTTCGCGATCTT
 CGCCGTTACCTTCTTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGAC
 AAGCTGCAGGAATTCAGGGATTACTCCAACGAAGAAAAAGATGGTAATCTTCCAGATATT
 GTGAATAGTGAAGTTTGCATGAGTTCCTGGTTAATTGTCATGAGAGATATGGGCCTGTGGT
 CTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTAAGCAGC
 ATATCAATCCCAATAAGACATCGGACCCCTTTGAAACCATGCTGAAGTCATTATTAAGGTAT
 CAATCTGGTGGTGGCAGTGTGAGTGAAAACCACATGAGGAAAAAATTGTATGAAAATGGTGT
 GACTGATTCTCTGAAGAGTAACTTTGCCCTCCTCCTAAAGCTTTGAGAAGAATTATTAGATA
 AATGGCTCTCCTACCCAGAGACCCAGCACGTGCCCTCAGCCAGCATATGCTTGGTTTTGCT
 ATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAAGATGATCAGGAAGTCATTCTG
 CTTCAGAGAATCATGGCACAGTTTGGTCTGAGATTGAAAAAGGCTTTCTAGATGGGTACAC
 TTGATAAAAAACATGACTCGGAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTT
 TTAAGGAACATCATAAAGAACGAAAAGGAAGGAAGTTCAGTCAACATATTTTCATTGACTC
 CTTAGTACAAGGGAACCTTAATGACCAACAGATCCTAGAAGACAGTATGATATTTCTCTGG
 CCAGTTGCATAATAACTGCAAAATTGTGTACCTGGGCAATCTGTTTTTAAACCACCTCTGAA
 GAAGTTCAAAAAAATTATATGAAGAGATAAACCAAGTTTTTGGAATGGTCTGTACTCC
 AGAGAAAATTGAGCAGCTCAGATATTGTGAGCATGTGCTTTTGTGAAACTGTTGGAAGTCC
 AACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATT
 CCTAGAGAGACCTCGTCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCC
 ATCTCCACACAAGTTTGATCCAGATCGGTTTGATGATGAATTAGTAATGAAAACCTTTTCCT
 CACTTGGATTCTCAGGCACACAGGAGTGTCCAGAGTTGAGTTTGCATATATGGTGACCACA
 GTACTTCTTAGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAGGGACAGGTATTGA
 AACAAAGTATGAAGTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGAT
 ATTAAATTTTATACATTTAAATCATTGTTAAATTGATTGAGGAAAAACAACCATTTAAAAA
 AAATCTATGTTGAATCCTTTTATAAACAGTATCACTTTGTAATATAAACACCTATTTGTAC
 TTAA

FIGURE 118

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVNSGSLHEFLVNLH
ERYGPVVSFWFGRRLVVSGLTVDVLKQHINPNKTSDPFETMLKSLRYQSGGGSVSENHMRK
KLYENGVTDSLKSNFALLLLKSEELLDKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFE
DDQEVIRFQKNHGTWVSEIGKGFLDGS�DKNMTRKKQYEDALMQLESVLRNIIKERKGRNFS
QHIFIDSLVQGNDLNDQQILEDSMIFSLASCIITAKLCTWAICFLTTSSEEVQKKLYEEINQVF
GNGPVTPEKIEQLRYCQHVLCE TVRTAKLTPVSAQLQDIEGKIDRFII PRETLVLVYALGVVL
QDPNTWPSPHKFD PDRFDELVMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLS
VEGQVIETKYELVTSSREEAWITVSKRY

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 271-290

FIGURE 119

CTAGATTTGTCGGCTTGC GGGGAGACTTCAGGAGTCGCTGTCTCTGAACCTCCAGCCTCAGA
GACCGCCGCCCTTGTCCCGAGGGCC**ATG**GGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTC
CTGACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACAT
ACAGGCCTGCCTGCCTCTCACGTTCACCCCGAGGAGTATGACAAGCAGGACATTACAGCTGG
TGGCCGCGCTCTCTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGA
GTCTCCATGTTCAACAGACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGT
GGCCCTGTCCTTCTTCATATTCGAGCGTTGGGAGTGCAC TACGTATTGGTACATTTTGTCT
TCTGCAGTGCCCTTCCAGCTGTCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAA
AAGAAACCCTTCT**TGA**ATTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGG
CCGCTTCGTATTCTTGAAGAAGGAAGGCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTC
TGCTGGAGGATATGTGTTGGAATAATTACGCTCTTGAGTCTGGGATTATCCGCATTGTATTTA
GTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACA
ATTAAAAAAAAAAAA

FIGURE 120

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACPLTFTPEEYDKQDIQLVAALSVTLG
LFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVT
EMALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

121/249

FIGURE 121

TCCCGGACCCCTGCCGCCCTGCCACTATGTCCCGCCGCTCTATGCTGCTTGCCCTGGGCTCTCC
CCAGCCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATA
GTGCCCCGAACGAGTGGAAGGCCCTGGCATCAGAGTGCGCCAGCACCTGAGCCTGCCCTT
ACGCTATGTGGTGGTATCGCACACGGCGGGCAGCAGCTGCAACACCCCGCCTCGTGCCAGC
AGCAGGCCCGGAATGTGCAGCACTACCACATGAAGACACTGGGCTGGTGCGACGTGGGCTAC
AACTTCCTGATTGGAGAAGACGGGCTCGTATACGAGGGCCGTGGCTGGAACCTTCACGGGTGC
CCACTCAGGTCACTTATGGAACCCCATGTCCATTGGCATCAGCTTCATGGGCAACTACATGG
ATCGGGTGCCACACCCAGGCCATCCGGGCAGCCAGGGTCTACTGGCCTGCGGTGTGGCT
CAGGGAGCCCTGAGGTCCAACATATGTGCTCAAAGGACACCGGGATGTGCAGCGTACACTCTC
TCCAGGCAACCAGCTCTACCACCTCATCCAGAATTGGCCACACTACCGCTCCCCCTGAGGCC
CTGCTGATCCGCACCCCATTCCTCCCCTCCCATGGCCAAAAACCCACTGTCTCCTTCTCCA
ATAAAGATGTAGCTC

FIGURE 122

MSRRSMLLAWALPSLLRLGAAQETEDPACCSPIVPRNEWKALASECAQHLSLPLRYVVVSHT
AGSSCNTPASCQQQARNVQHYHMKTLGWCDVGYNFLIGEDGLVYEGRGWNFTGAHSGHLWNP
MSIGISFMGNYMDRVPTPQAIRAAQGLLACGVAQGALRSNYVLKGHRDVQRTLSPGNQLYHL
IQNWPHYRSP

Signal peptide:

amino acids 1-20

123/249

FIGURE 123

CTGGACCCCGGAAAAAGAGAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTG
GACTCGCTGCTGCTTCTGTGTTCTCTGGTGCGAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGG
ATGATGGTCTCTCCGGCTCAGAGGACCTTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCC
CGGGTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCT
CCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCC
CGAACACAGCCCCCACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTAC
TCCAACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAA
TGGGACCTTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCC
TCGTGCCCCCCAGTAAAGCTGTAGAGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAG
GCCTCCA AAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCGCCGAGCCTC
GCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCAGCCACCT
GAGAGTCTCCACAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTAT
CGGCTGGTCCAGAGGTGTGCCAGATTACAACATACCATAGTGATACCCCTACTACCCATC
TGGGTGAGCCCGGGGCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGC
AGGAGACCATCTGGACACCGGGCAGGGAAGGGGTTGGCGCTCAGGCAGGAGGGGGGTGGAG
ACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTGAAGTGGCAGAGAAAGGGTCCCAAGTG
CTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACCTGGAGGAGGATGGGCT
CTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGGCCACAGAGAGATGCTGGGTCCCGAGGGCC
TGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGGTT
CTTGCCATCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTATCAGTGTGGACAGCCTG
TCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGG
CAGAGGAGCTCTCAGCCCTGCCTAGTGGGCGCCCTGAGCCCTTGTCTGTGCTGAGCATG
GCATGAGGCTGAAGTGGCAACCTTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTC
CAGCCAGGCCACCCCTTTCCA AAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATT
GCTGATGGCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCCACTAAGGCC
ACAGCCCATCCGGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAG
CATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACGGGGTTCTCC
CGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGA
GCATGTGCTGGATCTGTTCTGTGTCTGTCTGTGGTGGGGGGAGGGGAGGGAACTTGT
GAAACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTCTTGAGACAGGAAATAAAGCTT
GCCCCGGGGCA

FIGURE 124

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521
><subunit 1 of 1, 252 aa, 1 stop
><MW: 28127, pI: 8.91, NX(S/T): 5
MQLTRCCFVFLVQGSLLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMAN
STLLGLLAPPGAEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVD
HGNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGR
RTSLCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVCPTYNYHSDTPY
YPSG
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
21

GTGAATGTGAGGGTTTGATGACTTTAGATGATGTCTAGGAACACAGAGTGGGTGACAGGGCCCCCA
GCGAGGGCTGATTTCTTGGGCGGAGGAGTAGGGTAAAGGGTTCTGCATGAGCTCCTTAAAG
GACAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGACTTTGACTTCAAGCCACAGAAT
TGGTGAAGTGTGCGCGCCGCCGCCGCTCGCTCCTGCAGCGCTGTGACCTAGCCGCTAG
CATCTTCCCAGCACCGGGATCCCGGGTAGGAGGCGACGCGGGCGAGCACCAGCGCCAGCC
GGCTGCGGCTGCCACACGGCTCACCATGGGCTCCGGGCGCCGGGCGCTGTCCGCGGTGCC
GCCGTGCTGCTGTTCTCACGCTGCCGGGGTGCCTGCTGGGCACAGAACGACACGGAGCC
CATCGTGCTGGAGGGCAAGTGTCTGGTGGTGTGCGACTCGAACC CGGCCACGGACTTCAAAG
GCTCCTCTTCTCCCCGTGCGGATATCGGTCCGGGCGGCCAACTCCAAGTGTGCTTCTCG
GCGGTGCGGAGCACCAACCACGAGCCATCCGAGATGAGCAACAAGACGCGCATTTACTT
CGATCAGATCCTGGTGAATGTGGTAATTTTTTTCATTGGAGTCTGTCTTTGTAGCACCA
GAAAAGGAATTTACAGTTTTCAGTTTTTTCAGTGATTAAAGTCTACCAGAGCCAAATATCCAG
GTTAACTTGATGTTAAATGGAACCAGTAATATCTGCCTTTGCGGGGGACAAAGATGTTAC
TCGTGAAGCTGACACGAATGGTGTCTGCTCTACCTAGATAAAGAGGATAAGGTTTACCTAA
ACGTGAGGAAAGGTAAATTTGGTTGGAGCTGGCAGTATTCACGTTTTTCTGGCTTTCTGGTG
TTCCCCCTATAGGATTCATTTCTCATGATGTTCAATCCAGGTGAGGGATGACCCACTCCTG
AGTTATTGGAAGATCATTTTTTCATCATGGATGTGATGCTTTTATTGGTTTCTCATGGGTG
GATATGGATTCTAAGGATTCTAGCCTGTCTGAACCAATACAAAATTTACAGATTATTTTGTG
TGTGTCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTA
ACAGTCAAAAGCTGTCTGCAAGACTTATTCTGAATTTCAATTTCTGGGATTACTGAATTAGT
TACAGATGTGGAATTTTATTGTTTAGTTTTTAAAGACTGGCAACCAGGTCTAAGGATTAGA
AAACTCTAAGTTCTGACTTCAATCAACGGTTAGTGTGATACTGCCAAAGAACTGTATACTG
TGTTAATATATTGATTATATTTGTTTTTATCTTTGGAATTAGTTTGTTTGGTTCTTGTA
AAAACCTGGATTTTTTTTTTTCAGTAAGTGGTATTATGTTTTCTCTTAAATAAGGTAATGAA
TGGCTTGCCACAAATTTACCTTGACTACGATATCATCGACATGACTTCTCTCAAAAAA
GAATGCTTCATAGTTGATTTTAAATGATATGTGAAGAGTCATATTTTCCAAGTTATATT
TTCTAAGAAGAAGATAGATCATAAATCTGACAAGGAAAAAGTTGCTTACCCAAATCTAAG
TGCTCAATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAAATCTTATACTTTATTGC
TCAACTTAATTAATGATTGATAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTT
TCCGTAGACATGACCACTTTATTAAGTGGTGGGATGCTGTTGTTCTAATTATACCTAT
TTTTCAAGGCTTCTGTTGATTTGAAGTATCATCTGGTTTTGCCTTAACCTTTAAATTTGTA
TATATTTATCTGTTTAGCTAATATTAATTCAAATATCCCATATCTAAATTTAGTGAATAT
CTTGTCTTTTGTAGCTCATGATGAATTCATAAAATATTTATGTCTGTATAGAATAAAGA
TTAATATATGTTAAAAAA

FIGURE 126

MGSGRRALSAVPAVLLVLTLPGLPVWAQNDTEPIVLEGGKCLVVCDSNPATDSKGSSSSSPLGI
SVRAANSKVAFSAVRSTNHEPSEMSNKTRIIFYDQILVNVGNFFTTLESV FVAPRKG IYSFSF
HVIK VYQSQT IQVNLMLNGKPVISAFAGDKDVTREAA TNGVLLYLDKEDKVYLKLEKGNLVG
GWQYSTFSGFLVFPL

Signal peptide:

amino acids 1-27

127/249

FIGURE 127

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTC
GCCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCTCATCGCCGGAGC
TTTCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTGTTTGGTTCATGGCAAGAGTCATTA
TTGACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTC
TATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTT
GAAGAGTATAAACCAGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCT
TGGGCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCCTATCTGACTCCTTGGGG
CCAGGCACAGTGGGCATTTCATGGAGATTCTCCTCAATTCCTTTATTCAGCTTTTCATGAC
GCTGGTCATTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTGTATGGCTGTGAGAAGA
AAAAGTGGGGCATCCTCCTTATCGTTCCTGACCCACCTGCTGGTGTGAGCCAGACCTTC
ATAAGTTCCTATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGTCTATGGGCAC
CTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAG
ACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGATTAACCTCAGGGAACCAGCACTTCCCAA
ACCGCAGACTACATCTTTAGAGGAAGCACAACCTGTGCCTTTTTCTGAAAATCCCTTTTCTG
GTGGAATTGAGAAAGAAATAAACTATGCAGATA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 128

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658
><subunit 1 of 1, 257 aa, 1 stop
><MW: 28472, pI: 9.33, NX(S/T): 0
MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDN
KDGPQTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGF
GIMSGVFSFVNTLSDSLPGPTVGIHGDSPOFFLYSAFMTLVIIILHVFVGIVFFDGCCKKKW
GILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKCLLCQDKN
FLLYNQSR
```

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

**Glycosaminoglycan attachment site.**

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

[illegible]

FIGURE 130

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66659
><subunit 1 of 1, 832 aa, 1 stop
><MW: 94454, pI: 6.94, NX(S/T): 12
MFALGLPFLVLLVASVESHLGVLGPKNVSQKDAEFERTYVDEVNSELVNIYTFNHTVTRNRT
EGVRVSVNVLNKQKGAPLLFVVRQKEAVVSFQVPLILRGMFQRKYLYQKVERTLCQPPTKNE
SEIQFFYVDVSTLSPVNTTYQLRVSRMDDFVLRTEGEQFSFNTTAAQPQYFKYEFPEGVDSVI
VKVTSNKAFFPCSVISIQDVLCPVYDLNNVAFIGMYQTMTKKAAITVQRKDFPSNSFYVVVV
VKTEDQACGGS LPPFYPAEDEFPVDQGHQRKTL SVLVSQAVTSEAYVSGMLFCLGIFLSFYLL
TVLLACWENWRQKKKTL LVAIDRACPE SGHPRVLADSFPGSSPYEGYNYGSFENVSGSTDGL
VDSAGTGDLSYGYQGRSFEPVGTRPRVDSMSSVEEDDYDTLT DIDSDKNVIRTKQYLYVADL
ARKDKRVLKKYQIYFWNIIATIAVFYALPVVQLVITYQT VVNVGTGNQDICYNFLCAHPLGN
LSAFNNILSNLGYILLGLLFLLIILQREINHNRRALLRNDLCALECGIPKHGFLFYAMGTALM
MEGLLSACYHVC PN YTNFQFDT SFMYMIAGLCMLKYQKRHPDINASAYSAYACLAIVIFFS
VLGVVFGKNTAFWIVFSIIHI IATLLSTQLYYMGRWKLD SGIFRRILHVLYTDCIRQCSG
PLYVDRMVLLVMGNVINWSLAAYGLIMRPNDFASYLLAIGICNLLLYFAFYIIMKLRSGERI
KLIPLLCIVCTSVVWGFALFFFQGLSTWQKTPAESREHNRCILLDFDDHDIWHFLSSIA
MFGSFLVLLTLDDDLDTVQRDKIYVF
```

Important features of the protein:**Signal peptide:**

amino acids 1-18

Transmembrane domains:

amino acids 292-317, 451-470, 501-520, 607-627, 751-770

Leucine zipper pattern.

amino acids 497-518

N-glycosylation sites.

amino acids 27-30, 54-57, 60-63, 123-126, 141-144, 165-168, 364-367, 476-479, 496-499, 572-575, 603-606, 699-702

FIGURE 131

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGC
 TCTCTTCTTGTGTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCGTG
 ACCATGGTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCAAGGC
 CCAGCCTGCAGAGCTGTCTGTGGAAGTTCAGAAAACTATGGTGGAAATTTCCCTTTATACC
 TGACCAAGTTGCCGCTGCCCGTGAGGGGGCTGAAGGCCAGATCGTGTGTAGGGGACTCA
 GGCAGGGCAACTGAGGGCCATTTGCTATGGATCCAGATTCTGGCTTCCGTGGTGGACCA
 GGCCCTGGACCGAGAGGAGCAGGCAGAGTACCAGCTACAGGTCAACCTGGAGATGCAGGATG
 GACATGTCTTGTGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTG
 CCCCATTCTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCC
 CTTCTCTTCCCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACCTCGGATCTTCGAT
 TCCACATCCTGAGCCAGGCTCCAGCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGG
 CTGGGGCTCTGGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGAC
 CTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTG
 CCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGTCCCTAGAGCCTATCCACCTGGCA
 GAGAATCTCAAAGTCTTATACCCGCACCATATGGCCAGGTACACTGGAGTGGGGTGATGT
 GCACATACCTCAGGAGAGCCATCCCCGGGACCCCTTTGAAGTGAATGCAGAGGGGAACCTCT
 ACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAG
 AATTCCCATGGCCAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTGTGATGGATGAGAA
 TGACAACCTGCCTATCTGCCCTCCCCGTGACCCACAGTCAAGTCCCTGAGTCTCAGTCCAC
 CAGTACTGAAGTGACTAGACTGTGACGAGAGGATGCAGATGCCCCCGGCTCCCCCAATTTCC
 CACGTTGTGATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCCTTCCA
 GGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCATCTCCGACAGGCCAGAAACA
 TCCTGCTTCTGGTGTGCGCATGGACCTGGCAGGCCGAGAGGGTGGCTTCAGCAGACGTGT
 GAAGTCAAGTCGAGTCAAGATATCAATGATCAGCCCTGAGTTCACTCACTTCCCAGAT
 TGGGCCTATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGTCAACAGCCA
 TTGATGCTGACCTCGAGCCCCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACA
 GAAGGGACTTTTGGCCTGGATTGGGAGCCAGACTCTGGGCATGTGTAGACTCAGACTCTGCAA
 GAACCTCAGTTATGAGGCAGCTCCAAGTCATGAGGTGGTGGTGGTGCAGAGTGTGGCGA
 AGCTGGTGGGGCCAGGCCAGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGA
 GTGATGCCACCCCCAAGTTGGACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCC
 AGCCGGCTCTTTTCCCTGCTGACCATCCAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCT
 CCCTAGTCAATGACTCAGAGGGCTGGCTCTGCATTGAGAAATTTCTCCGGGGAGGTGCACACC
 GCCCAGTCCCTGCAGGGCGCCAGCCTGGGGACACCTACACGGTGCTTGTGGAGGGCCAGGA
 TACAGCCCTGACTCTTGGCCCTGTGCCCTCCCAATACCTCTGCACACCCGCCAAGACCATG
 GCTTGATCGTGAGTGGACCCAGCAAGGACCCGATCTGGCCAGTGGGACAGGTCCCTACAGC
 TTCACCTTGGTCCCACCCACCGTGCACCGGATGGCGCCTCCAGACTCTCAATGGTTTC
 CCATGCCCTACCTACCTTGGCCCTGCATTGGTGGAGCCAGTGAACACATAATCCCCGTGG
 TGGTCAGGCCAATGCCCCAGATGTGGCAGCTCCTGGTTCGAGTGATCGTGTGCTCGCTGCAAC
 GTGGAGGGCAGTGCATGCCCAAGGTGGGCCGATGAAGGGCATGCCACAGAGCTGTGCGGC
 AGTGGGCATCCTTGTAGGCCCTGGTAGCAATAGGAATCTTCTCATCTCACTTTTCAACCC
 ACTGGACCATGTCAAGGAAGAAGGACCCGATCAACCAGCAGACAGCGTGGCCCTGAAGGCG
 ACTGTCTGAATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCCTGTGGCTCCATCTGAGTCCC
 CTGGAGAGAGCCACGACCCCAAGATCCAGCAGGGGACAGGACAGATGAGAAGCCCTCCAT
 CTGCCCTGGGTGGAGGCACCATCACCATCACCAGGCATGCTGCAGAGCTGGACACCCAA
 CTTATGGACTGCCCATGGGAGTGCTCCAAATGTGAGGGTGTTTGCCCAATAATAAAGGCCCA
 GAGAACCTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 132

MVPAWLWLLCVSVQALPKAQAELSVVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSG
 KATEGPFAMDPDSGFLLVTRALDREEQAQYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVP
 HFSQAIYRARLSRGRTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSMDMFQLEPRL
 GALALSPKGSTSLDHALERTYQLLVQVKMDQASGHQATATVEVSIESTWVSLEPIHLAE
 NLKVLYPHHMAQVHWSSGGDVHYHLESHPPGPFVEVNAEGNLYVTRELDREAQAEYLLQVRAQN
 SHGEDYAAPLELHVLMVDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSH
 VVYQLLSPEPEDGVEGRAQVVDPTSGSVTLGVPLPRAQONILLVLAMDLAAGGGFSSTCE
 VEVAVTDINDHAPEFITSQIGPISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTE
 GTFGLDWEPPDSGHVRLRLCKNLSYEAAPSHEVVVVQSVAKLVGPGPGPGATATVTVLVERV
 MPPPKLDQESYEASVPISAPAGSFLLTIQPSDPISTRLRFSLVNDSEGWLCIEKFSGEVHTA
 QSLQGAQPGDTYTVLVEAQDTALTAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSF
 TLGPNPTVQRDWRLQTLNGSHAYLTLALHWVEPREHII PVVVSHNAQMWQLLVRIVCRCNV
 EGQCMRKVGRMKGMPTKLSAVGILVGTIVAIGIFLILIFTHWTMSRKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 133

CCGGGGAC**ATG**AGGTGGATACGTTCATTGGGGCCCTTATTGGGTCCAGCATCTGTGGCCAA
GAAAAATTTTTTGGGGACCAAGTTTTGAGGATTAATGTCAGAAATGGAGACGAGATACGCAA
ATTGAGTCAACTAGTGAATTCAAACAACCTGAAGCTCAATTTCTGGAAATCTCCCTGCTCCT
TCAATCGGCCCTGTGGATGTCCTGGTCCCATCTGTCAGTCTGCAGGCATTTAAATCCTTCTCTG
AGATCCCAGGGCTTAGAGTACGCAGTGACAATTGAGGACCTGCAGGCCCTTTTAGACAATGA
AGATGATGAAATCAACACAAATGAAGGGCAAGAACGGAGCAGTAATAACTTCAACTACGGGG
CTTACCATTCCCTGGAAGCTATTATTACCACGAGATGGACAACATTGCCGAGACTTTCTTGAC
CTGGCGAGGAGGGTGAAGATTGGACATTCGTTTGAAAACCGGCCGATGTATGTACTGAAGTT
CAGCACTGGGAAAGCGGTGAGGCGGCCGCCGCTTTGGCTGAATGCAGGCATCCATTTCCGAG
AGTGGATCTCCCAGGCCACTGCAATCTGGACGGCAAGGAAGATTGTATCTGATTACCAGAGG
GATCCAGCTATCACCTCCATCTTGGAGAAATGGATATTTTCTTGTGCTGTGGCCAATCC
TGATGGATATGTGTATACTCAAACCAAACCGATTATTGGAGGAAGACGCGGTCCCCAAATC
CTGGAAGCTCCTGCATTTGGTGCTGACCCAAATAGAACTGGAACGCTAGTTTTGCAGGAAAG
GGAGCCAGCGACAACCCCTTGCTCCGAAGTGATACCATGGACCCACGCCAATTCGGAAGTGGA
GGTGAATCAGTGGTAGATTTTCATCAAAAACATGGGAATTTCAAGGGCTTCATCGACCTGC
ACAGCTACTCGCAGCTGCTGATGTATCCATATGGGTACTCAGTCAAAAAGGCCCCAGATGCC
GAGGAAGCTCGACAAGGTGGCGAGGCTTGCGCCCAAAGCTCTGGCTTCTGTGTGCGGGCACTGA
GTACCAAGTGGGTCCCCACTGCACCACTGTCTATCCAGCTAGCGGGAGGAGCATCGACTGGG
CGTATGACAAAGGCATCAAAATTTGCATTACATTTGAGTTGAGAGATACCGGAGCATGTAGGC
TTCTCTCTGCCAGCTAACCAGATCATCCCCACTGCAGAGGAGACGTGGCTGGGGCTGAAGAC
CATCATGGAGCATGTGCGGGACAACTCTAC**TAG**CGCATGGCTCTGCTCTGTCTACATTTAT
TTGTACCCACACGTGCACGCACCTGAGGCCATTGTTAAAGGAGCTCTTTCTTACTCTGTGTAG
TCAGAGCCCTCTGGGTTTGTGGAGCACACAGGCCTGCCCTCTCCAGCCAGCTCCCTGGAGT
CGTGTGCTCTGGCGGTGTCCTTGCAAGAACTGGTCTGCCAGCCTGCTCAATTTTGGTCTGT
CTGTTTTTGTATGAGCCTTTTGTCTGTTTCTCCTCCACCCTGTGGCTGGGCGGCTGCACCT
AGCATCACCCCTCTCTGGGTGGCATGTCTCTCTTACCTCATTTTAGAACCAAAGAACATC
TGAGATGATTCTCTACCTCATCCACATCTAGCCAAGCCAGTGACCTTGTCTGTGGGCACCT
GTGGGAGACCACTTGTCTTTAGGTGGGTCTCAAAGATGATGTAGAATTTCCCTTTAATTTT
TCGAGTCTTCTCGGAAAAATATTTTCTTTGAGCAGCAAATCTTTAGGGATATCAGTGAAG
GTCTCTCCCTCCCTCCTCCTGTTTTTTTTTTTTTTTGGAGACAGATTTTGTCTTTGTGTGCC
CAGGCTGGAGTGTGATGGCTCGATCTTGGCTCACCACAACCTCTGCCTCTGGGTTCAAGCA
ATTCTCCTGCTCAGCTCTTGTAGTGTGTTTATAGGCGCATGCCACCATGCTTGGCTA
ATTTTGTGTTTTTGTAGTAGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTCAAACCTCCA
ACCTCAGGTGATCTGCCCTCCTTGGCCTCCCAGAGTGCTGGGATTACAGGTGTGAGCCACTG
TGCGGGGCGCGTCCCTCTTTTATAGGCTGAATACAAAGTAGAAGATCACTTTGCTTCTTAC
TGTGCTGAGAATTTCTAGATACTACAGTTCTTACTCCTCTCTTCCCTTTGTTATTACGTGTG
ACCAGGATGGCGGGAGGGGATCTGTGTCACTGTAGGTACTGTGCCAGGAAGGCTGGGTGAA
GTGACCATCAAAATGAGGATGGTGAATTTATCCCCATCTGTCTTAATGGGCTTACCTCCT
CTTTGCCCTTTGAACCTCACTTCAAAGATCTAGGCCCTCATCTTACAGGTCCATAACTCACTAT
CTGGCTGATAATCTCACTGCCCTGGCACATTTCCCATTTGTGTGTGTGTGTGTGTGTGTGTGT
TCCTTGTCTCTGGTTTGT
TCTGTCTATTTTGTATCTCGACCAAGTTCCTAAGTAGAGCAAGAATTCATCAACCAGCT
GCCCTTGTGTTTCAATTTCACTCTCAGCACGTACCATCTGCTCTTTGTGTGTGTGTGTGTGT
TTGTTTTTGTCTTTTACCAACATGTCTGTAAATCTTAACCTCCTGCCTAGGATTTGTACA
GCATCTGGTGTGTGCTTTATAAGCCAATAAATATTCAATGTGAAAAAAAAAAAAAAAAA

1015571-121101

FIGURE 134

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKNFWKSPSSFNR
PVDVLVPSVSLQAFKSFLRSQGLEAYVTIEDLQALLDNEDDEMQHNEGQERSNNFNFGAYH
SLEAIYHEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRPAAVWLNAGIHSREWI
SQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANPDGYVYTQTQNRLWRKTRSRNPGS
SCIGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKSVVDFIQKHGNFKGFIDLHSY
SQLLMYPYGYSVKKAPDAEELDKVARLAALKALASVSGTEYQVGPTCTTVYPASGSSIDWAYD
NGIKFAFTFELRDTGTYGFLLPANQIIPTAETWLGLKTIMEHVRDNL

Signal peptide:

amino acids 1-16

10015671-12340

[illegible]

FIGURE 136

MASYLYGVLF~~FAVGL~~CAP~~IYCV~~SPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLVLE
 TPSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVP
 SKDLTLKMG~~SALFV~~KKELQLQANFLGNV~~KRLYEAE~~VSTDFSNPSIAQARINSHVKKKTQGK
 VVDIIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAF
 GVDTELNCFVLQMDYKGD~~AVAFFVL~~PSK~~GKMRQ~~LEQALSARTLIKWSHSLQKRWIEVFI~~PRF~~
 SISASYNLETILPKMGIQNAFDKNADFSGI~~AKRDS~~LQVSKATHKAVLDVSEEGTEATAATTT
 KFIVRSKDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

Signal peptide:

amino acids 1-20



FIGURE 138

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATI
SGSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNESSESTTSSGASTATNSE
SSTPSSGASTVTNSGSSVTSSGASTATNESSESTVSSRASTATNESSTLSSGASTATNSDSS
TTSSGASTATNESSESTTSSGASTATNESSESTVSSRASTATNESSESTTSSGASTATNESRTT
SNGAGTATNESSESTTSSGASTATNSDSSVTSSGASTATNESSESTTSSGASTATNESSESTTSS
GASTATNSDSSSTTSSGAGTATNESSESTVSSGISTVTNSESTPSSGANTATNESSESTTSSGA
NTATNESSESTVSSGASTATNESSESTTSSGVSTATNESSESTTSSGASTATNSDSSSTTSSEAST
ATNESSESTVSSGISTVTNSESESTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAV
SEAKPGGSLVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGP
GPGGNHGAPHRPRWSPNWFWRPVS SIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 139

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTC
CCAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCC
CTGTCTGGATGGGCGGCCAGCGATGACCCATTGAGAAGGTCATTGAAGGGATCAACCGAGG
GCTGAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGAATCACGC
ATGCCGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGC
AAGGAGTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGAT
CAACCATGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACG
CTGCTGGACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCACACTGGGGTCCAC
CAGGCTGGGAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGG
AAAGGAAGTGGAGAAGCTTGGCCAAGGTGCCCACCATGCTGCTGGCCAGGCCGGGAAGGAGC
TGCAGAATGCTCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGC
AACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGG
GGCCTCAGTCAACACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCA
TGCCCT**TAA**ACTGGCATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTACATCAGCTGAC
ATGACCTGGAGGGGTTGGGGTGGGGACAGGTTTCTGAAATCCCTGAAGGGGTTGTACTG
GGATTTGTGAATAAACTTGATACACCA

FIGURE 140

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675

><subunit 1 of 1, 247 aa, 1 stop

><MW: 25335, pI: 7.00, NX(S/T): 0

MHLARLVGSCSLLLLLGGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAG
 REVEKVFNGLSNMGSHTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHVNNAAAG
 QAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQN
 AHNGVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP

Important features of the protein:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

TCGGGTTCCCGACGGGCTTCGCGCGGGCGGCGCTGGCAAGGGGACGAGTCAGTGGACATCCAGGAAGACGGCG
CCGCGGGGGGCGATGACGCTGGCTGACCTGACTCACTACGAGTCCGGAGCGGGGCGCCCGGGGCGACTCG
GGGCGGGACCGCGGGGGCGAGCTCGCCGCCGCTGAGTCGGCGCGAGCCATCTGAGCCGACGCGCGGGAGACCGCTG
CTCTCTGCTGTTCTCGGAATCTGGCAGCGGATGGGCTCGAGAGCTGCTGCCGCCCTATGGGCGCGCTGCGC
CTCGGCAACCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCAAGCGCGCGCTCCGAGCTGGGCGCTCAGC
CCCCGATGACGCTGCTCTGGCTGCTGAGAAAGGCGCATTCAGATTGAAAGCTGACACATCTCTCCACTAC
ACAGCGCTTTCTGCTGACGAGGATGSGCAGGACCTGTGTCGTGGTGCTCGAGAGGCTCTTTTGCACTGATGAG
AACTCTCAGCTTCTCGGACGGGGGATCCAGGAGAGCTGCTTTGGGGTGGAGACGAGGAAGAACAAGCACTG
AGCTTCAAGGGCAGGAGCCACGAGCGGACTGTCAAATCATCAAGATCTCTCGCGCTCAGCGGAGCTCAG
AGTGTTCACCTGTGGCAGCAGCGCTTCAGCCCGATGTGACTCATCAACATGGAGAACTTCAAGCTGGGAAG
CAGCAGAAAGGGAATGTCTCTGGAAAGTGGCAGGCGGCTGTTCCTTCGACCAGAAATTCACGCTCACTGCTG
TGTTGGTGTGATGGGCACTCTACACTGGAACAGTCAGCAGTTTCAAGAGGAATGACCGCGCATCTCGCGGAG
CAAAGCGTTCGGCCGACCAAGACGAGAGTCCCTCACTGGCTCAAGGCGACCGAGTTTGTGGCTCAGCGTAC
ATTCTGAGAGCGCTGGGACCTTTCGAGGCGATGATGACAAGATCTACTTTTCTCAGGAGAGCTGGCCAGGAA
TTTGAGTTTGTGAGAAACCAATTTGTCCGCATTGCGCCGATTCGAAAGGGCGATGAGGGTGGAGAGGGGTG
CTACAGCAGCGCTGGACCTCTCTCTCAGGCGCAGCTGTGTGCTGACGCCCGGACAGCTTGTGGCTTCCCTTCAAC
GTGTCTCAGGAGTTCTTTCAGCTGAGCGCAGCCCGCCAGGATGGCTGGACACCTTTCTATGGGCTTCTACT
TCCGATGGGACAGGGGAATCAAGAGGCTCTGCGCTGTGTGCTTCAAAATGAAGATGTGCAGAGAGTTCTT
AGCGCGCTCTACAAGGAGTGAACCTGGAGACACGAGTGGTACCGTGAACCCCGGTGGCCACACCGCGG
CTCGGAGCGCTCAATCCAAACAGGCGGGGAAGGAAGATCAACTCATCTCTGACGCTCCGACAGCGCGCTGCTG
AACTTCTCTGAAGACCACTTCTGATGGAGCGGAGCTCGAAGCGCGCATGCTGCTGCGAGCCCAAGCTCG
TACGAGCGCGCTGAGGCTATACCGCGTCTCTGGCTGACACACACTACCTGACGTCTCTCTCTGGGCATGTGTG
GGCCGGCTCCAGAGGCTCAGCGGTGGCGCCCGGGTGCACATCAATGAGAGTCTGCAGTCTTCTCATCGGA
CAGCGCCCTGCAGAACTCTGCTCTGGACACCCAGGGGGCTGCTGTATCGGGCTACACTCGGGCGTAGTCAG
GTGCCCATGGCCAACTGCAGCTGTACCGGAGCTGTGGGATGCTCTCTGCGCCGGAGCCCTCACTGTGCTGG
CGCGCTCCAGTGTGAAGCAGTGAAGCTCTACCCAGCTCAGTGGCCACAGCGCTGATCATCAGGAACTGAG
GGAGCAGCGCGCAAGAGGCTTTCAGCGGCTCTTGCTTTGTGCTCCGCTTTTGTGACCAAGGGGAGAAGCCA
CTGTGAGCAAGTCAGATTCCAGCCCAACACAGTGAACATCTTGGCTTGGCCCTCTCTCTCAACCTGGGACCGCA
TCTGGAGCTACGAAACGGGGGCCCCGCTCAATGCTCGGCTCTGTCAGCATGCTCACTCAGTGGGACCTGTGCTG
GTGGGCGCCCAACAGCTGGGGAGTTCGATGCTGTGCTACGAGAGGGGCTCCAGACAGCTGTGATAGGAGCTAC
TGCCAGAGGTGTGTGAGAGCAGGGGTGGCAGACAAACAGATGAGGTTGGCAGTGTACCGTCAATTCATGACCA
TCGCGTGTGATGTCACAGCTGTGTGGCAGGGCCAGCTGGGTTGCAGACAGCTCTCAATGGAAGAGTTCCTGTG
ATGTGACGCTCTTTTGTGCGCGCTGCTGCTCCAGTTTATTTCTTCTGCTACCGGACCGGAACAGCATGAA
GTCTTCTCGAAGCAGGGGAATGTGCAGCGCTGCACGCCAAGAGCTGCCCTGTGGTGTGCCCTCTGAGACCGCC
CACTCAACCGCTAGGCGCCCTTAGCCACCTGCTCGATACAGCGAGGGTACAGTCCCTGTAGACAGCCCCCG
GGGGCCCGAGTCTTCTAGTACGATGAGAGAGGACCACTCAGCATCCAGACAGCACTGTGGAGGATCTCCCAAGT
TGCCCGCCCGGCTCGCGCTTGGCTGGAGATCGTGATCTGTGTGTGTGAGAGCTCATCCGAGAGAGCTG
TGCCCTGGCTCAGGGGCTGTGAATGCTCGGAGAGGGTCAACTGGAACCTCCCTCGCTCTGCTCTTGTGGAAC
ACAGCGCTGTGTGCGGGCCCTTGGAGCTCTGGAGCGCTGGCGCTGCTGCTGCTCAAGTCAAGTAGCAGGACTCC
TACCCACCAAGACACCCAAACAGCGTGGCCCGAGAGTCTCGGCCAAATATGGGGGCTGCTGATTTGTGTGAA
CAGTCCCTCTATGTGATACTGAGCCCTTTGTTTATAAAACAAATCAAATGTGAACTAGATAGTGGGGGAAG
ATAGATGATGCTGACGACACAGCGCTGCTCAGTTCTATGGCTCCGAGGGGTGCTGGGATGAGTCTCAAAGTGT
TTGCTGAGACAGATGTGGAACCTCCACCACTGGCTCTTCACTTCCCAATATGTCCTGCTGCCACGGGTGCT
CTGTCTCTCAGTTCAGATTCAGGACAGCTTGGGCTGGCTGCGTTTGCCTTGCAGCTCAGCGAGGATGTATGTTG
TTGTGCTGCTCTGTCACCAACTCAGGACACGAGGCTAGTTGGCATGCGGCGCTCACAGGCTCTGGGCTG
GGACCCAACTCTGGACCTTTCCAGCTGTATCAGGCTGTGGGCACAGAGAGGACAGCGCAGTCAAGAGATA
TTTGTGCAAGTGTACGCTTCTTCCCTCAGAAATTCAGGGAAGAGCATGTGCGCTGCTCTCGCGTGTGTGCTGTA
GAAACCGCTGTGCGCCTTCCCAACCTATCCACCTCTGCTCTCACTTTTGAATCAACACAGGAACTAATGCACTG
CTGTGCTGCTCCCGAGTCCCGAGTCAACCTCCATCCCTCACTTCTCACTCAAGGATATCAACACTGCC
AGCACAGGGGCCCTGAATTTATGTGGTTTATTATACATTTTAAATAGATGCATTTATGTCAATTTTTTAAATAA
GTCTGAGAAATTAAGTTTATAAAATAAAAAA

FIGURE 142

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962

><subunit 1 of 1, 837 aa, 1 stop

><MW: 92750, pI: 7.04, NX(S/T): 6

MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLQPPPTWALSPRISLPLGSEERPFLRF
 EAEHISNYTALLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEKKQQCSFKG
 KDPQRDCQNYIKILLPLSGSHLFTCGTAAFSPMCTYINMENFTLARDEKGNVLLEDGKGRCP
 FDPNFKSTALVVDGELYTGTVSSSFQGNDFPAISRSQSLRPKTESLNLWLQDPAFVASAYIPE
 SLGSLQGDDDKIYFFFFSETQGEFEFFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLC
 SRPDDGFFPNVLQDVFTLSPSPQDWRDTLFYGVFTSQWHRGTTEGSAVCVFTMKDVQRVFSG
 LYKEVNRETQQWYTVTHPVPTPRPGACITNSARERKINSSLQLPDRVLNFLKDHFLMDGQVR
 SRMLLLQPQARYQRVAVHRVPGHLHTYDVLFLGTGDGRHLKAVSVGPRVHII EELQIFSSGQ
 PVQNLLDTHRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSlyQP
 QLATRPMWQDIEGASAKDLCSASSVVSFSFVPTGEKPCEQVQFPNTVNTLACPLLSNLATR
 LWLRNGAPVNASASCHVLPTGDLLLVGTQQLGEFQCWSLEEGFQQLVASCYCEVVEDGVADQ
 TDEGGSVPVPIISTSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSM
 KVFLKQGECA SVHPKTCPVVLPPETRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTSEKR
 PLSIQDSFVEVSPVCPRPVRVLGSEIRDSVV

Transmembrane domains:

amino acids 23-46 (type II), 718-738

FIGURE 143A

10015671212110

CTAAGCCGGAGGATGTGCAGCTGCGCGCGGGCGCGGCTACGAAGAGGACGGGGACAGGCGCGCTGCGAACCGGA
 GCCCAGCCAGCGCGGAGGACGCGGGCAGGGCGGGGACGGGAGCCCGGACTCGCTGCGCCGCGCGCTGCTGCGCGGTG
 TCGCGGCCCGCGCTCCCCGGCGGAGCGGGAGGAGCGCCGCCACCTCGCGCGGACGCGCGCGCTGACGCGCGCG
 CCGGAGTATGGTCCCTCTTAAGGCGCGAGCGCGCGCGCGCGGGCGGGTGTGCGGAAACAAGCGCGCGCGCGGGG
 CTTGCGGCGGCGCTCGGGGCGCGGATGGCGCGCGCGGGCGCGCGCGCGCGCGCTGCGCCGGCGCGGGCTCG
 CGCGCGTATGGCGGGGCTGCGCTCGTGGCGGGGCGAGCGGCTGAGGGCGCGCGGAGCGCTGCGCGCGCGCGCGG
 GCGCGCGCGCGCGCGCGCGGGCGGAGCGCGCGGGC**ATG**CGCGCGCGCGCGCGCGCGCTGCGCTGAGCTGCTG
 TGCCTCGGGCTCGTCTGGGCTTCTGCTGCTGCGGCTCGGGCTCGTCTGCCCGGGCTTCCGAGCTGAAGCGAGCGG
 GCCCAGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGTCCGGCGAGGGCGCGGCTTCCAGCGCGCGCGCGCGCGG
 GCGATGCGCGCGGGCGCAGCTTGGCGCGCGCGCTCGGACCCAGATGGCGCGCGCGCGCAGCAGAACATTTCTCT
 TCGTGGGAGTCA TGACCGCCAGAAATACCTG CAGATCGGGCGCTGGCGCGCTACAGAACATGTGTTCAAGACAA
 TTCTCTGGGAAGTTCAGTTCTTCTCAAGTGAGGGTTCTGACACATCTGTACCAATTCAGTAGTGCCACTACGGG
 GTGTGGACGACTCTACCCGCCCCAGAGAAGTCTTCTCATGTGCTCAAGTACATGCACGACCACTACTGGACA
 AGTATGAATGGTTATGAGAGCAGATGATGACGTGTACATCAAAGGAGACCGTCTGGAGAATTCTCTGAGGAGTT
 TGAACAGCAGCGAGCCCTCTTCTTGGGCAGACAGGCTGGGCACACGGAAGAATGGGAACACTGGCCCTGG
 AGCCTGGTGAGAATTCTGATGGGGGGCTGGCGTGATCATGAGCGGGAGTCTCTCGAGAATTGTCGCGG
 ACATTGGCAAGTGTCTCGGGAGATGTACACCACCATGAGAGCTGGAGGCTGTGTCGGAGGTTTG
 CAGGGTGCAAGTGTCTGCTGCTGTATGAGATCGCGAGCTTTTTTATGAGAATTAAGGAGCTGTGTCGGAGGTTG
 ACATTAGAGATTCCTATACGTAAATTCACCAAGACTATCACTACACCCCAACAAACACCCCTCCAGCT
 ACAGGCTCCACAGCTACATGCTGAGCGCAGAATATCCGAGCTCGGCCATTCGCAATACAGGCTGCGCCGGGAAA
 TTTCTCTGATGACGAATAACGACACACAGAATTCATAAAGAGGACCTCCAGCTGGGAATCCCTCCCTCTTCA
 TGAAGTTTCAGCCCCCGCAGCGAGGAGGAGATTCTGGAAATGGGAGTTTCTGACTGGAATAATCTGTTCTGGCAG
 TTGACGCGCCAGCCCTCGAGAGGGAATGGACTCGCGCCAGAGGGGAGGCTTGGACGACATTTGTCATCAGGCTG
 TCGAGATGATCAATGCCAACGCCAGACAGCAGGGCGCATCATTGACTTCAAAGAGATTCAGTAGGCTACGGCTACCGC
 GGGTGAACCCCATGTATGGGCTGAGTACATCTTGGACCTGCTGCTTCTGACAAAAAGCAGAAAGGAGAAAA
 TGACGGTCCCTGTGAGGAGGCGACGCTATTTACAGAGACTTTCAGCAAAATCAAGTTTGTGGAGCATGAGGAGT
 TGGATGCACAAGAGTTGGCCAAAGAGAATCAATCAGGAATCTGGATCCTGTCTCTTCTCTCAAACTCCCTCGAAGA
 AGCTCGTCCCTTTACGCTCGCTGGGTGGAAGAGTGAGCAACAAGAACCAAAGATAAAAAGATAAACATCATGA
 TTCTCTTGTCTGGGCGTTTTCGACATGTTTGTGAGATTTATGGGAACTTTGAGAAGAGCTGTCTTATCCCCAATC
 AGAAGCTCAAGCTCGGTTCTGCTTTTCAATTTGACTCCAAACCTGACAAGGCCAAACAAGTTGAACATGATGA
 GAGATTACCGCATTAAGTACCTTAAGCCGACATGACAGATTTTGCCTGTGCTCTGGAGATTTTCAAGAGCGCTGG
 CCTTGGAAAGTAGGATCTCCCACTTTAAACAATGAATCTTTGCTCTTCTCTCGCAGCTCGACTCGCTGTTTACTA
 CAGAATTCCTTCAGCGATGTCGAGCAAAACAGTTCTGGGCCAAACAATATATTTTCCAATCATCTTCAGCGAGT
 ATGACCCAAAGATTGTTATAGTGGGAAAGTTCCCACTGACAACCACTTTGCCCTTACTCAGAAAAGTCGGCTCT
 GGAAGAACTATGGGTTGGCATCAGCTGATTTATAAGGGAGATCTTGTCCGAGTGGGTGGCTTTGATGTTTCCA
 TCCAAGGCTGGGGCTGGAGGATGTGGACCTTTTCAACAGGTTTGTCCAGGAGGTTTGAAGACGTTTAGGAGCC
 AGGAAGTAGGAGTAGTCCAGCTCCACCATCTCTTTTGTATCCCAATCTTGACCCCAACAGTACAAAATGT
 GCTTGGGTCCAAAGCATCGACTTGGGTCCACCCAGCAGCTGGCTGAGATGTGGCTGGAAAAAAATGATCCAA
 GTTACAGTAAAGCAGCAATAATAATGGCTCAGTGAGGACAGCC**TAAT**GTCCAGCTTGTCTGGAAAAAGCGTTT
 TAATTAATCTCAATTTATTTTCAAAAATTTTGTATGATCAGTTTGAAGTCGATACAGGATATATTATTC
 AAGTGGTTTCTTACATAGGACTCCTTTAAGATTGAGCTTCTGAACAAGAGGTTGATCAGTGTTTGCTTTGAA
 CACATCTCTCTGTGCAACATATGTAGCAGACCTGTTAACTTTGACTTGAATACCTTGTGATGAACAAAATCTT
 TTTAAAAAATGTTTCTTTTGAGACCGTTTGTCTCCAGTCTTGGCAGCAAAACGTGAACATCTCTGCAAAAGTAT
 TATTGTAAACAAAACATCTTAACCTGCTAAATGTTCTGTGTGATTGTAAACTCCACAGATTCTACCTTTTGT
 GTTTTGTTTTCTTTTACAATTTGTTTAAAGCCATTTTCATGTTCCAGTTGTAAAGTAAAGGAATGCGATAATA
 GCTGTTTTCATCATTGTCTCAGGAGGCTTCCAGAGTTGATCATTCTCTCATGTGACTCTGCTCAGCATGGC
 CAGCTAGGTTTGTGTTGTTTGTGTTTGTCTTTTGTGAGACGGAGTCTCACTCTGTTAACCGAGCTGGAAATG
 GCGGGCGCAATCTGGCTCACTTTAACTCCACTTCCCTGGTTCAGCAATTCCTCGCTTGTCTTCCCTCCAGAT
 AGCTGGGATTACAGGCACACACACACCGCCAGTAGTTTTTGTATTTTATGACAGCGGGGTTCACCAT
 GCAAGCCAGCTGGCCAGCTAGTTTTHAAGCAAGGGCGGTGAAGAAGCAGCATGAGATGTGGCTGTCTCTG
 TGGTAGTTTCTTCCGCTTAATAGACTGSCATTAAATTTCAAGAAGGATTGGCAATTTCTCTCTTGACCCCTT
 CTTCTTAAAGGGTAAAAATTAATGTTTGAATGACAAAGATGAATTAACAATAATCTGATGACACAGACT
 GAAACATGACCATACACCTTAATCAAACGTTGGGGAATAATGTATTTGGTTTGTCTTCTCATCTGTCTG
 TGTATCTGCGTGAGATGTTTTCATCTTTTATTACTGTTTGTGTTATCTTGTATCTGTATGAATAACCTTAA
 TTTATTTAATACTGTTGTTTCAGAGCTCGCCATTCTTGATACCTGTTAGTTAGTATTAATTTATGTTGATGAT
 GAGTGTGTTTGTGTTTATTTTTCAGTAACCCGATCTCCAAGATTTCCTTTTGGAAACGCTTTTCCCTTC

FIGURE 143B

TTAATTTTATATTCCTTACTGTTTACTAAATATTAAGTGTTCTTTGACAATTTTGGTGCTCATGTGTTTTGGG
GACAAAAGTGAAATGAATCTGTCATTATACCAGAAAAGTTAAATTCCTCAGATCAAATGTGCCTTAATAAATTTGTT
TTCATTTAGATTTCAACAGTGATAGACTTGCCATTTTAAATACAGTCATTGGAGGGCTGCGTATTTGTAAATAG
CCTGATGCTCATTTGGAAAAATAAACAGTGAACAATATTTTCTATTGTACTTTTCGAACCATTTTGTCTCATT
ATTCTGTTTTAGCTGAAGAATTGTATTACATTTGGAGAGTAAAAAACTTAAACACGAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 144

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836
><subunit 1 of 1, 802 aa, 1 stop
><MW: 91812, pI: 9.52, NX(S/T): 3
MAARGRRAWLSVLLGLVLGFVLASRLVLPRASELKRAGPRRRASPEGCRSGQAAASQAGGAR
GDARGAQLWPPGSDPDGGRDRNFLFVGVMTAQKYLQTRAVAAYRTWSKTIPGKVQFSSSEG
SDTSVPIPVVPLRGVDDSYPPQKKSFMMMLKYMHDHYLDKYEWFMRADDDVYIKGDRLNFLR
SLNSSEPLFLGQTGLGTTEEMGKLALEPGENFCMGGPGVIMSREVLRRMVPHIGKCLREMYT
THEDVEVGRCVRRFAGVQCWVSYEMRQLFYENYEQNKKGYYIRDLHNSKIHQAITLHPNKNPP
YQYRLHSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMRFPQPRQREE
ILEWEFLTGKYL SAVDGP PPRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQY
GYRRVNP MYGA EYILD LLLYKHKHGKKMTVPVRRHAYLQQTFSKIQFVEHEELDAQELAKR
INQESGSL SFLSNSLKKLVFPQLPGSKSEHKPEPKDKKINILIPLSGRFDMFVRFMGNFEKTC
LIPNQNVKLVLVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQ
FNNESLLFFCDVDLVFTTEFLQRCRANTVLGQQIYFPPIIFSQYDPKIVYSGKVPSDNHFAFT
QKTGFWRNYGFGITCIYKGD LVRVGGFDVSIQGWGLEDVDLFNKVVQAGLKTFRSQEVGVVH
VHHPVFCDPNLDPKQYKMCLGSKASTYGSTQQLAEMWLEKN DPSYSKSSNNNGSVRTA
```

Signal peptide:

amino acids 1-23

FIGURE 145

GGACAACCGTTGCTGGGTGTCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCC
 TTTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGT
 ACGGACGACGCC**TATGA**AGCCCTTAGTCCCTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTG
 TGCCGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAACTTGAATCATTATATA
 CAAGTTTATAGAGAACCTAGTACGAAGTGTCCCTCTGGGGAGCCAGGTCGTGAGAAAAATC
 TAACTCTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAATTTAAGGAGCTAGTTA
 CACATGGAGACGCTTCAA**CTGAGA**ATGATGTTTTAA**CCAATCCTATCAGTGAAGAACTACA**
 ACTTTCCCTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAACACACGGAAAGTACCCC
 ATTCTGGTCGATCAAACCAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTG
 AAAATGAAGAGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAACTGAGGCACCAAGAATG
 TTGCCAGTTGTTACTGAATCATCTACAAGTCCATATGTTACCTCATA**CAAGTCA**CCTGT**CA**C
 CACTTTAGATAAGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCCTCAGCTCT
 CAGGTGAACTGCGATAGAAAAACCCGAAGAGTTTGGAAGCACCCAGAGAGTTGGAATAAT
 GATGACATTTTGAAAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCAC**TTCTTAGTGA**
 CACCAGCAACCCAGCATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCC
 TTGCTCTAGCAGCAGCAGCAGAACATAAATTAAAAACAATGTATAAGTCCAGTTATTGCCA
 GTAGGACGAACAAGTAATAAAATTTGATGACATCGAACTGTTATTAACATGCTGTGTAATTC
 TAGATCTAAACTCTATGAATATTTAGATATTAATGTGTTCCACCAGAGATGAGAGAAAAAG
 CTGCTACAGTATTCAATACATTAAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTA
 AAAGTTTAT**TAA**ACAATAATATAAAATTTTAAACCTACTTGATATTCATAACAAAGCTGA
 TTTAAGCAAACTGCATTTTTTACAGGAGAAATAATCATATTCGTAAATTTCAAAGTTGTAT
 AAAAAATATTTCTATTGTAGTTCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACA
 ATTTTCATATGCCTAAAAACCTAATTTAAATAAAATTTTGGTTCAAGAAAAA

FIGURE 146

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68864

><subunit 1 of 1, 350 aa, 1 stop

><MW: 39003, pI: 5.59, NX(S/T): 1

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRVSPSGEPGREKKSNSPK
HVYSIASKGSKFKELVTHGDASTENDVLNPISEETTFPTGGFTPEIGKKKHTESTPFFWSI
KPNNVSIVLHAEOPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSTSPYVTSYKSPVTTLDK
STGIEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNP
AYREDIEASKDHLKRSALAAAAEHKLKTMYSQLLPVGRTSNKIDDIETVINMLCNSRSKL
YEYLDIKCVPPPEMREKAATVFNTLKNMCRSRRTALLKVY

Signal peptide:

amino acids 1-19

GGGCTCGAGCGGCTCTAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTG
ACCTCCAAATCATCCATCAACCCCTGCTGTGCATCTGTTTTCATATGATGTGAGATCAACCCACA
GGAATATCG**CATGGCT**TTTGTGCTCATTTTGTCTCTCAGTTTCTACGAGCTGGTGTGAGGACA
GTGGCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACCGCGTGTCTC
CCTGTCCTCTTTTCTCGAGACCAGTGCAGAGGTCATGGAAGTCGGGTTCTTCAGGAATCAG
TTCATGCTGTGGTCCACCTCTACAGACAATGGGGAGACTGGGAATCTAAGCAGATGCCACA
GTATCGAGGGAGAACTGAGATTTGTGAAGGACTCCATTGCGAGGGGGCGTGTCTCTTAAGG
TAAAAAATCATCTCCCTCGGACATCGGCTGTATGGGTGCTGTTTCAGTTCCGAGGATTTAC
GATGAGGAGGCCACTTGGAGCTCGGGGCTGGCAGCACTGGGCTCACTTCTCTCATTTCCAT
CGTGGGATATGTTGACGGAGGTATCCAGTTACTCTGCCGTCTCAGGCTGTCCCCAAGC
CCACGCCAAGTGGAAAGGTCACACAGGACAGGATTTGTCTCAGACTCCAGAGCAAAATGCA
GATGGGTACAGCCTGTATGATGTGAGATCTCCATTTATAGTCCAGGAAAATGTCTGGGAGCAT
ATTGTGTTCCATCCACCTGTGTCGAGCAGAGTCATGAGGTGGAATCCAAAGTATTGATAGGAG
AGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGT
GCCCTGTGGTGGTGTGTTGATCGGGATGATAATTTGTTTCTCAAATCCAAAGGGAAAATCCA
GGCGGAATCGGACTGGAGAAGAAAGCACGGACGGAGCAGAATGTGAGAGACGCCGGAAACGCA
CAGTGGAGGTGACTCTGTGATCCAGAGACGGCTACCCCGAAGCTTCGCGTTTCTGATCTGAAA
ACTGTAAACCCATGAAAAGACTCCCCAGGAGTGCTCACTCTGAGAAGAGATTTACAGGAA
GAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGAGGACATTACTGGGAGGTGGACGTGGGAC
AAATGTAGGGTGGTATGTGTTGGGAGTGTGTCGGGATGACGTAGA CAGGGGGGAAGAACATGTG
ACTTTGTCTCCCAACATGGTATTTGGGTCCTCAGACTGACAACAGAACTTTGATTTATCG
ATTCATATCCCCATTTTATCAGCCTCCCCCCAGACCCCTCTACACGAGTATGGGGTCTTCC
TGGCATATGAGGGTGGACCATCTCCTTCTTCAATACAAATGACAGGTCCCTATTATTATCC
CTGCTGCATCTGCAGTTTGAAGGCTTGTGTGAGACCCATATCCAGCATCGCATGTATGACGA
GGAAAGGGGACATCCCATATTATATGTGTCAGTGTCTCTGGGGT**TGAC**AGCAGAGAACCTGC
CTTAAAGGGCCCCACACCAAGACCCAGACAGCAGCAAGGGAGAGTGCTCCCCAGACGTTGC
CCCAGCTTCTCTCCGGAGCCTGCGCACAGAGAGTCAGCGCCCCACTCTCCTTTAGGGAGC
TGAGGTTCTCTCGCCCTGAGCCCTCGACAGCGGCGATCAGCCTTCCAGATGAGGGGGAT
TGCCCTGACCCCTGGGAGTCAGAAGCCATGGCTGCCCTGAAGTGGGGAGCGGAATAGACTCA
CATTAGGTTAGTTTGTGAAAACTCATCAGCTAAGGCACTTTGAACAGGTCACAACTCC
CAGGCTCCTCATTTGTCTAGTCAGGACAGTGATTCTTGCTCAGAGTGAAAGTAAAGAGA
CAACGAATGTGAATCATGCTTGCAGGTTTGAGGGGACAGTGTTTGCTAATGATGTGTTTTTA
TATTATACATTTTCCCCACAATAACTCTGTTTGCTATTCTACATTAATTTACTTTTCTGA
TACCAAATCACCATTGGAATGATTTATGAACACTGCTTTGTGAGGCTCAAGAGATAATACGTA
GAGGTAGGATTTTCTACTGATTTATTAAGCCACGATTAACCTGATACCAAACACGGCAAGG
AAACACAGAAGAAGAGGAAGGAAACACTACAGGTCATATCCCTCATTAAACACAGACAAAAA
TTCTAAATAAAAATTTTAAACAAATTAACATAACAAATATATTTAAAGATGATATATAACTACT
CAGTGTGGTTTTGTCCCACAAAATCGAGATGGTTTAAATTTAAATATCAACCGTGTAAT
CAGCACAATTAATAAGATGAAAAGAAACCAATTAATAAAAAAAAAAAAAA

FIGURE 148

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866
><subunit 1 of 1, 466 aa, 1 stop
><MW: 52279, pI: 6.16, NX(S/T): 2
MAFVLILVLVSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHA
VVHLYRDGEDWESKQMPQYRGRTFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEE
ATWELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGY
SLYDVEISIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSFPWRLASILLGLLCGALC
GVVMGMIIIVFFKSKGKIQAELDWRKKGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFTTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLS
PNNGYWVLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLT
CQFEGLLRPYIQHAMYDEEKGTPIFICPVSWG
```

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 149

CCTTCACAGGACTCTTCATTGCTGGTTGGCA**ATG**ATGTATCGGCCAGATGTGGTGAGGGCTA
 GGAAAGAGTTTGTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTC
 CTGGCAGTGTGCATTGGACTCACTGTTTCATTATGTGAGATATAATCAAAGAAGACCTACAA
 TTACTATAGCACATTGTCATTTACAACGACAACTATATGCTGAGTTTGGCAGAGAGGCTT
 CTAACAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAA
 TCTCCATTAAAGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGG
 AGTGTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAG
 ATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTA
 GATCCTCACTCAGTTAAAAATAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCA
 TTGCTGCGGAACACGAAGAAGTAAAACTTAGGTGAGAGTCTCAGGATCGTTGGTGGGACAG
 AAGTAGAAGAGGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCACTCGCTGT
 GGAGCAACCTTAATTAATGCCACATGGCTTGTGAGTGTGCTCACTGTTTTACAACATATA
 GAACCCCTGCCAGATGGACTGCTTCCTTTGGAGTAACAATAAAACCTTCGAAAATGAAACGGG
 GTCTCCGGAGAATAATTGTCCATGAAAAATACAAACCCCATCACATGACTATGATATTTCT
 CTGTCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGA
 TGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAA
 ATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAAC
 TGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTT
 AGAAGGAAAAACAGATGCATGCCAGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTA
 GAGATATCTGGTACCTTGCATGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCAACAG
 CCTGGTGTTTATACATAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAAGTGGTATCT**TA**
AGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTTGGGTGTGGAGGCCATT
 TTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAAC
 TGTTTGCTTGATGCATGTATTTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTG
 CCAGATCAACTCTGTCACTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATA
 ATACAATATTACATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTGTGAGAATTTTGAC
 TTGTTGACATAAATTTGTAATGCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTC
 AGCTCCTCTCATTTACGCAATATCCATTTTCAAGGTGCAGAACAGGAGTGAAAGAAAAATA
 TAAGAAGAAAAAATCCCCACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGT
 GGAATATTAGAAATGATCATATTTCATTATGAAGGTCAAGCAAGACAGCAGAATACCAATC
 ACTTCATCATTTAGGAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGAGCCAAAGATATA
 TCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTG
 ACCTATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATAT
 TTATTTAACATTGTTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 150

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871
```

```
><subunit 1 of 1, 423 aa, 1 stop
```

```
><MW: 47696, pI: 8.96, NX(S/T): 3
```

```
MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTINYISTLSFTTD
KLYAEFGREASNNFTEMSQRLESMVKNAFYKSPLEEFVKSQVIKFSQQKHGVLAHMLLICR
FHSTEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTL
GQSLRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFG
VTIKPSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSPVPYTNVHRVCLPDASYEFQPGDV
MFVTGFGALKNDGYSQNHRLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGD
SGGPLVSSDARDIWIYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI
```

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 151

GTCGAAGGTTATAAAAGCTTCCAGCCAAACGGCATTGAAGTTGAAGATACAACCTGACAGCA
 CAGCCTGAGATCTTGGGGATCCCTCAGCCTAACACCCACAGACGTCAGCTGGTGGATTCCCG
 CTGCATCAAGGCCTACCCACTGTCTCCATGCTGGGCTCTCCCTGCCTTCTGTGGCTCCTGGC
 CGTGACCTTCTTGGTTCCAGAGCTCAGCCCTTGGCCCTCAAGACTTTGAAGAAGAGGAGG
 CAGATGAGACTGAGACGGCGTGGCCGCTTTGCCGGCTGTCCCCTGCGACTACGACCACTGC
 CGACACCTGCAGGTGCCCTGCAAGGAGCTACAGAGGGTCGGGCGCGCGGCCTGCCTGTGCC
 AGGACTCTCCAGCCCCGCCAGCCGCCGACCCGCCGCGCATGGGAGAAGTGCGCATTGCGG
 CCGAAGAGGGCCGCGCAGTGGTCCACTGGTGTGCCCCCTTCTCCCCGGTCCCTCCACTACTGG
 CTGCTGCTTTGGGACGGCAGCGAGGCTGCGCAGAAGGGGCCCCCGCTGAACGCTACGGTCCG
 CAGAGCCGAAGTGAAGGGGCTGAAGCCAGGGGGCATTATGTCGTTTGCCTAGTGGCCGCTA
 ACGAGGCCGGGGCAAGCCGCGTGCCCCAGGCTGGAGGAGAGGGCCTCGAGGGGGCCGACATC
 CCTGCCTTCGGGCCTTGACGCCGCTTGCGGTGCCGCCAACCCCGCACTCTGGTCCACGC
 GGCGCTCGGGTGGGCACGGCCCTGGCCCTGCTAAGCTGTGCCGCCCTGGTGTGGCACTTCT
 GCCTGCGCATCGCTGGGGCTGCCCGCGCCGAGCCGCCGCCGAGCCGCAGGGGCGCTCTGA
 AAGGGGCTGGGGGCATCTCGGGCACAGACAGCCCACTGGGGCGCTCAGCCTGGCCCCG
 GGAAGAGAGAAAACCGCTGCCTCCAGGGAGGGCTGGACGGCAGCTGGGAGCCAGCCCCAG
 GCTCCAGGGCCACGGCGAGTCATGGTTCTCAGGACTGAGCGCTTGTTTAAAGTCCGGTACTT
 GCGCTTTGTTTCTGCTGAGGTCTGGGAAGGAATAGAAAGGGGCCCCCAATTTTTTTTTTA
 AGCGGCCAGATAATAAATAATGTAACCTTTGCGGTTAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2

FIGURE 152

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68874
><subunit 1 of 1, 238 aa, 1 stop
><MW: 25262, pI: 6.44, NX(S/T): 1
MLGSPCLLWLLAVTFLVPRAQPLAPQDFEEEEADETETAWPPLPAVPCDYDHCRLQVPCKE
LQRVGPAACLCPGLSSPAQPDPDPPRMGEVRIAAEEGRAVVHWCAPFSPVLHYWLLWLDGSEA
AQKGPPLNATVRRRELKGLKPGGIYVVCVVAANEAGASRPQAGGEGLEGADIPAFGPCSRL
AVPPNPRTLVAHAVGVGTALALLSCAALVWHFCLRDRWGCPRRAAARAAGAL
```

Important features of the protein:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 194-220

N-glycosylation site.

amino acids 132-135

FIGURE 153

AGAGAAAGAACGCTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCC
 CTGCCCGCATGAGCCCCCGCGTGCCTCCCCGACTATCCCCAGGCCGGGCGTGGGGCACCAGGG
 CCCAGCGCCGACGATCGCTGCCGTTTGGCCCTTGGGAGTAGGATGTGGTGAAGGATGGGGC
 TTCTCCCTTACGGGGCTCAC**AATGG**CCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCT
 CTACGCCCTCAATCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGA
 TGAGGGACTACCTAAATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTC
 ATTTTGACTACTTTCTCTGTGTTTCATCCGGTCATGATTGCTGTTTGGCTTTCTTATCAT
 TGTGGGATGTTAGGATATTGTGGAACGGTGAAGAAATCTGTTGCTTCTTGCATGGTACT
 TTGGAAGTTTGGTTGTCATTTCTGTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAG
 GAACCTATGGTCCAGTACAATGGTCAGATATGGTCACTTTGAAAGCCAGGATGACAAATTA
 TGGATTACCTAGATATCGGTGGCTTACTCATGCTTGGAAATTTTTTCAGAGAGAGTTTAAAT
 GCTGTGGAGTAGTATATTTCACTGACTGGTGGAAATGACAGAGATGGACTGGCCCCAGAT
 TCCTGCTGTGTAGAGAATCCCAGGATGTTCCAAACAGGCCACCAGGAAGATCTCAGTGA
 CCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTTTTGGAGGAACCAACAACCTGC
 AGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAATCTTGCCATGATTCTCACC
 ATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGGACAGACCAATGATGTC
 CTTGAAGAATGACAACTCTCAGCACCTGTCACTGCTCCCTCAGTAGAACTGTTGAAACCAAGCC
 TGTCAAGAATCTTTGAACACACATCCATGGCAACAGCTTTAATACACACTTTGAGATGGAG
 GAGTT**TAA**AAAGAAATGTACAGAAAGAAAACCAACAACCTGTTTTATTGGACTTGTGAATT
 TTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAAATGTTGCCATAAAATAACA
 CCTAAGCATATACTATTTCTATGCTTTAAATGAGGATGGAAGTTTTCATGTATAGTCAC
 CACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGC
 CTGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTT
 CGCATCCATGCCAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTT
 CTACCAACTAGTATATAAAGTACTAAATAAATGCTAACATAGGAAGTAGAAAATACATAATA
 ACTTTTATTACTCAGCGATCTATTTCTCTGATGCTAAATAAATATATATCAGAAAACCTTC
 AATATTGGTGACTACCTAAATGTGATTTTGTCTGTTACTAAATATTCTTCCACTTAAAA
 GAGCAAGCTAACACATTGCTTAAAGCTGATCAGGGATTTTTTGTATATAAGCTGTGTATAA
 TCTGTATAAATTCAGTCGATTTTCAGTTCGTATAATGTTAAGAAATAACCATATGAAAGGAAA
 ATTTGCTCTGTATAGCATCATTTATTTTAGCCTTTCTGTTAATAAAGCTTTACTATTCTGT
 CCTGGGCTTATATTACACATATAACTGTTATTTAAATACTTAACCACATAATTTGAAAATTA
 CCAGTGTGATACATAGGAATCATTATTCAGAATGTAGTCTGGTCTTTAGGAAGTATTAATA
 GAAAATTTGCACATAACTTAGTTGATTGAGAAAGGACTTGTGATGCTGTTTTCTCCCAATG
 AAGACTCTTTTGCACATAAAGCTTTTAAAAAGCTTATCTTTGCCCTTCTCCAACAAGAA
 GCAATAGTCTCCAAGTCAATATAAATTTCTACAGAAAATAGTGTTCTTTTCTCCAGAAAAT
 GCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATTCTTTGTTTATTTCACTGATTA
 ATATACTGTGGCAAATTCACAGATTATTAATTTTTTTTACAAGAGTATAGTATATTTATTT
 GAAATGGGAAAAGTCAATTTTACTGTATTTTGTGTTATTTGTTTATTTCTCAGAATATGGAA
 AGAAAATTAATGTGTCAATAAATATTTTCTAGAGAGTAA

1015671.12101

FIGURE 154

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880

><subunit 1 of 1, 305 aa, 1 stop

><MW: 35383, pI: 5.99, NX(S/T): 0

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPV
VHPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQ
WSDMVTLLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREF
PGCSKQAHQEDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWAL
YYDRREPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 155

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCC
 TGCACCTCGGGCCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGA
 CCTGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAG
 ACCGGGAGGATCAGAGGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGC
 CTCGATGTCAAACCCCTGCGCAAACCCCTATCCCCATGGAGACCTTCAGAAAGGTGGGGAT
 CCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTGTCTCATCAAGG
 TGATCTGTGATAAATACTACTTCCCTCTGCGGGCAGCCTCTCCACTTCATCCCAGGAAGCAG
 CTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACGTGTGTCAAGAGCTT
 CCCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGG
 ACTCGGCCACAGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAG
 ACAGCCTGTAGGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGGCCAGACCAGGATCTGGA
 TGTGTTGAAATCACAGAAAAAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTCT
 TCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCCTGTGGGAAGAGCCTGAAGACCCCCGT
 GTGGTGGGTGGGGAGGAGGCCCTCTGTGGATTCTTGGCCTTGCCAGGTGAGCATCCAGTACGA
 CAAACAGCACGTCTGTGGAGGGAGCATCCTGGAACCCCACTGGGTCTCACGGCAGCCCACT
 GCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGC
 AGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATATTGAATTCAACCCCATGTACCCCAA
 AGACAATGACATCGCCCTCATGAAGCTGCAGTTCCTCACTTTCTCAGGCACAGTCAGGC
 CCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGA
 TGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCA
 GGTCAATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGAAGTACCAGAGAAGA
 TGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCC
 CTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGG
 GGGCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATG
 TCTGGAAGGCTGAGCT**GTA**TGCTGCTGCCCCTTTGAGTGTCTGGGAGCCGCTTCCTTCTCTG
 CCCTGCCCACCTGGGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACC
 CTCTGCCACAGCCTCAGCATTTCTTGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACC
 TCGCAGCCAGAGGCGCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCC
 AGCATCCAGGGGAGAGACACAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAAGAA
 GGAACCTTCCCACTACTGAATGGAAGCAGGCTGTCTGTAAAGGCCAGATCACTGTGGG
 CTGGAGAGAGAAGGAAAGGCTCTGCGCCAGCCCTGTCCGTCTTCACCCATCCCCAAGCCTA
 CTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTT
 ACCTACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCT
 CTGGCAAAAAAAAAA

FIGURE 156

```

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47644, pI: 5.18, NX(S/T): 2
MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVVLIKVILDKYYF
LCGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAAVRLSKDRSTLQVLDSATGNWF
SACFDNFTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGCPCLSGSLVSL
HCLACGKSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCVGSILDPHWVLTAAHCFRKHTDV
FNWKVRAGSDKLGSFPPLAVAKIIIEFNPMYPKDNIDIALMKLQFPLTFSGTVRPICLPFFD
EELTPATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPE
GGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

```

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 158

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71166

>subunit 1 of 1, 761 aa, 1 stop

>MW: 83574, pI: 6.78, NX(S/T): 4

MALPALGLDPWSLLGLFLFQLQLLLPTTTAGGGGQGPMPRVRYAGDERRALSFFHQKGLQ
DFDTLLLSGDBGNTLYVGAREAILALDIQDPGVPRLKNMIPWPASDRKKSECAFKKKSNETQC
FNFIRVLVSYNVTHLYTCGTFAFSPACTFIELQDSYLLPISEDKVMEGKGQSPFPDAHKHTA
VLVDGMLYSGTMNNFLGSEPILMRTLGSQPVLKTDNFLRWLHHDSFVAAIPSTQVYFFFE
ETASEFDFFERLHTSRVARVCKNDVGGKLLQKKWTTFLKAQLLCTQPGQLPFNVIRHAVLL
PADSPTAPHIYAVFTSQWQVGGTRSSAVCAFSLLDIERVFKGKYKELNKETSRWTTYRGPET
NPRPGSCSVGPSSDKALTFMKDHFLMDEQVVGTPLLVKSGVEYTRLAVETAQGLDGHSHLVM
YLGTTTGSLHKAVVSGDSSAHLVEEIQLPDPEPVRNLQLAPTQGAVFVGFGSGVWRVPRAN
CSVYESCVDCVLARDPHCAWDPESTRCCLLSAPNLNSWKQDMERGNPEWACASGPMRSRLRP
QSRPQIIKEVLAVPNSILELPCPHLSALASYIWSHGPAAVPEASSTVYNGSLLLIQDGVGG
LYQCWATENGFSYPVISYWVDSQDQTLALDPELAGIPREHVKVPLTRVSGGAALAAQQSYWP
HFVTVTVLFALVLSGALIILVASPLRALRARGKVQGCETLRPGEKAPLSREQHLQSPKECRT
SASDVDADNNCLGTEVA

Signal peptide:

amino acids 1-30

Transmembrane domains:

amino acids 136-156, 222-247, 474-490, 685-704

FIGURE 159

AGGGTCCCTTAGCCGGGCGCAGGGCGCGCAGCCAGGCTGAGATCCGCGGCTTCCGTAGAAG
 TGAGC**ATG**GCTGGGCAGCGAGTGTCTTCTCTAGTGGGCTTCCCTTCCCTGGGGTCCGTGCTC
 TCAGAGGCTGCCAAAATCCTGACAATATCTACAGTAGGTGGAAGCCATTATCTACTGATGGA
 CCGGGTTTCTCAGATTCTTCAAGATCACGGTCATAATGTCAACCATGCTTAACCCAAAAAGAG
 GTCCTTTTATGCCAGATTTTAAAAAGGAAGAAAAATCATATCAAGTTATCAGTTGGCTTGCA
 CCTGAAGATCATCAAAGAGAATTTAAAAAGAGTTTGTATTCTTTCTGGAAGAACTTTAGG
 TGGCAGAGGAAAAATTTGAAAACCTATTAAATGTTCTAGAATACTTGGCGTTGCAGTGCAGTC
 ATTTTTTAAATAGAAAGGATATCATGGATTCTCTTAAAGAAATGAGAACTTCGACATGGTGATA
 GTTGAAACTTTTGACTACTGTCCCTTCTGATTGCTGAGAAGCTTGGGAAGCCATTGTGGGC
 CATTCTTTCCACTTCATTGGCTCTTTGGAATTGGGCTACCAATCCCCTTGTCTTATGTTTC
 CAGTATTCGGTTCCCTTGCTGACTGATCACATGGACTTCTGGGGCCGAGTGAAGAATTTCTG
 ATGTTCTTTAGTTTCTGCAGGAGGCAACAGCACATGCAGTCTACATTTGACAACACCATCAA
 GGAACATTTACAGAAGGCTCTAGGCCAGTTTTGTCTCATCTTCTACTGAAAGCAGAGTTGT
 GGTTCATTAACCTCTGACTTTTGCCCTTGATTTTGTCTCGACCTCTGCTTCCCAACACTGTTTAT
 GTTGGAGGCTTGATGGA AAAACCTATTAAACCAAGTACCACAAGACTTGGAGA ACTTCATTGC
 CAAGTTTGGGGACTCTGGTTTTGTCTTGTGACCTTGGGCTCCATGGTGAACACCTGTGAGA
 ATCCGGAAATCTTCAAGGAGATGAACAATGCCCTTTGCTCACCTACCCCAAGGGGTGATATGG
 AAGTGTCAAGTGTCTCATTTGGCCCAAGATGTCCACCTGGCTGCAATGTGAAATTTGTGGA
 CTGGCTTCTCAGAGTGACCTCCTGGCTCACCCAAGCATCCGTCTGTTTGTCAACCCACGGGC
 GGCAGAATAGCATAATGGAGGCCATCCAGCATGGTGTGCCCATGGTGGGGATCCCTCTCTTT
 GGAGACCAGCCTGAAAACATGGTCCGAGTAGAAGCCAAAAAGTTTGGTGTTCCTATTACAGTT
 AAAAGAAGCTCAAGGCAGAGACATTGGCTCTTAAGATGAAACAAATCATGGAAGACAAGAGAT
 ACAAGTCCGCGGCAGTGGCTGCCAGTGTCTCCTGCGCTCCCACCCGCTCAGCCCCACACAG
 CGGCTGGTGGGCTGGATTGACCACGTCTCCAGACAGGGGGGCGGACGCACCTCAAGCCCTA
 TGTCTTTAGCAGCCCTGGCATGAGCAGTACCTGTTGACGTTTTTGTGTTTCTGCTGGGGC
 TCACTCTGGGGACTCTATGGCTTTGTGGGAAGCTGCTGGGCATGGCTGTCTGGTGGCTGCGT
 GGGGCCAGAAAGGTGAAGGAGACAT**TAA**GGCCAGGTGCAGCCTTGGCGGGGTCTGTTTGGTGG
 GCGATGTCAACATTCTTAGGGAGCTTCCCACTAGTTCTGGCAGCCCCATTCTCTAGTCCTTC
 TAGTTATCTCCTGTTTTCTTGAAGAACAGGAAAAATGGCCAAAAATCATCCTTTCCACTTGC
 TAATTTTTGCTACAAATTCATCCTTACTAGCTCCTGCCTGCTAGCAGAAATCTTTCCAGTCTC
 CTTGTCCTCCTTTTGTGGCCATCAGCAAGGGCTATGCTGTGATTCTGTCTCTGAGTGACTTG
 GACCACTGACCCCTCAGATTTCAGCCCTTAAATCCACCTTCCCTCTCATGCGCCTCTCCGAA
 TCACACCCCTGACTCTTCCAGCCTCCATGTCCAGACCTAGTCAGCCTCTCTCACTCCTGCCCC
 TACTATCTATCATGGAATAACATCCAAGAAAGACACCTTGCAATATCTTTACAGTTTCTGTTTT
 TGTTCTCCACATATTTCTCTTCAATGCTCAGGAAGCCTGCGCTTGAGAGTTCAAGGCG
 CGGACACAGGCTCACAGGTCTCCACATTTGGGTCCCTGTCTCTGGTGCCACAGTGAGCTCCT
 TCTTGGCTGAGCAGGCATGGAGACTGTAGGTTTCCAGATTTCTTGAAAAATAAAAGTTTACA
 GCGTTATCTCTCCCAACCTCACTAA

FIGURE 160

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71169
><subunit 1 of 1, 523 aa, 1 stop
><MW: 59581, pI: 8.68, NX(S/T): 1
MAGQRVLLLVGFLLPGVLLSEAAKILTISTVGGSHYLLMDRVSQILQDHGHNVTMLNHKRGF
FMPDFKKEEKSYQVISWLAPEDHQREFKKSFDFFLEETLGGRGKFENLLNVLEYLALQCSHF
LNRKDIMDSLKNENFDMVIVETFDYCPFLIAEKLGGKPFVAILSTSFGSLEFGLPIPLSYVPV
FRSLITDHMDFWGRVKNFLMFFSFCRRQQHMQSTFDNTIKEHFTGSRPVLSHLLLKAELWF
INSDFAFDFARPLLNTVYVGGLEKPIKPVPPQDLENFIAKFGDSGFVLVTLGSMVNTCQNP
EIFKEMNNAFAHLPPQGVWKQCQSHWPKDVHLLAANVKIVDWLPQSDLLAHPSIRLFVTHGGQ
NSIMEAIQHGVPVMVGIPLFGDQPENMVRVEAKKFGVSIQLKKLKAETLALKMKQIMEDKRYK
SAAVAASVILRSHPLSPTQRLVGWIDHVLQTGGATHLKPYVFQQPWHEQYLFDFVFVFLGLT
LGTWLWLCGKLLGMAVWWLRGARKVKET
```

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 483-504

FIGURE 161

GGGCTGTTGATTGTTGGGGGATTTGAAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCT
GCCTCTGGGCATATGCACACACTCACACATCTCTGTACACCCGCTCACACACATACCATGTT
CTCCATCCCCCAGGTCAGCCCTCAGTGCTGTCCCATCCAGCAGGGCTACCCCTGAAGCTCT
GGCTGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCCTTTCTCTCCCAAAGCCCA
ACTGCTGTCACTGCATGCTCTGCCAAGGAGGAGGGAAGTGCAGTGACAGCAGGAGTAAGAGT
GGGAGGCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTCAGCGAGCCCTAGAGAGGGC
AGACTATCAGGGTGCCGGCGGTGAGAATCCAGGAGAGGAGCGGAACAGAAAGAGGGGCAGA
AGACCGGGGCACTTGTGGGTTGCAGAGCCCTCAGCCATGTTGGGAGCCAAGCCAACTGGC
TACCAGGTCCCCACACAGTCCCGGGCTGCCCTTGGTTCTGGTGCTCTGGCCCTGGGGGCC
GGGTGGGCCCAGGAGGGGTGAGAGCCCGTCTGCTGGAGGGGGAGTGCTTGGTGGTCTGTGA
GCCTGGCCGAGCTGCTGCAGGGGGGCCCGGGGAGCAGCCCTGGGAGAGGCACCCCTGGGC
GAGTGGCATTGTGCTGGTCCGAAGCCACCACATGAGCCAGCAGGGGAAACCGGCAATGGC
ACCAAGTGGGGCCATCTACTTCGACAGGTCCTGGTGAACGAGGGCGGTGGCTTTGACCGGGC
CTCTGGCTCCTTCTGAGCCCTGTCCGGGGTGCTACAGCTTCGGTTCCATGTGGTGAAGG
TGTACAACCCGCAAACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCTCTCAGCC
TTTGCCAATGATCCTGACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGA
CCCTGGGGACCGAGTGCTCTGCGCTGCGTCGGGGGAATCTACTGGGTGGTTGGAATACT
CAAGTTTCTCTGACCTTCTCTATCTTCCCTCTCTGAGGAGCCCAAGTCTTTCAAGCACAAGAAT
CCAGCCCCGTGCAACTTTCTTCTGCCCTCTCTTGCCCCAGAAACAGCAGAGGCAGGAGAG
ACTCCCTCTGGCTCCTATCCACCTCTTTGCTGGGACCTGTGCCAAACACCCAAGTTTAA
GAGAAGAGTAGAGCTGTGGCATCTCCAGACCAGGCCCTTCCACCCACCCACCCCAAGTTACC
CTCCAGCCACCTGCTGCATCTGTTCTGCTGCGCTGAGGATCAGGGCAAGGTTTGGA
AGAAGGAAGATCTGCACTACTTTGCGGCTCTGCTCCTCCGGTCCCCACCCCAAGTCTCT
GCTCAATGCTGATCAGGGACAGGTGGCGCAGGTGAGCCTGACAGGCCCCACAGGAGCCCAG
ATGGACAAGCCTCAGCGTACCCCTGCAGGCTTCTTCTGTGAGGAAGCCAGCATCACGGATC
TCAGCCAGCACCGTCAGAAGCTGAGCCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCAC
AGGCAGAAGGGTGGGAAGGGCTGGAGTCTGTGGTGGTGAAGGAAGGAGGGGTGATTG
TCTAGACTGAACATGGTACACATTCTGCATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCT
GGCTGTCTTCTATGCTGGATCCCAGATGGACTCTGGCCCTTACCTCCCCACCTGAGATTAG
GGTGAGTGTGTTTGCTCTGGCTGAGAGCAGAGCTGAGAGCAGGTATACAGAGCTGGAAGTGG
ACCATGGAACATCATGAACCATGCATCCTCTTGTCTGGCCACCTCTGAAACTGCTCCAC
CTTTGAAGTTTGAACCTTAGTCCCTCCACACTCTGACTGCTGCTCCTTCTCTCCAGCTCTC
TCACTGAGTTATCTTCACTGTACCTGTTCCAGCATATCCCCACTATCTCTCTTCTCCTGAT
CTGTGCTGTCTTATCTTCTCTTAGGCTTCTATTACCTGGGATTCATGATTCTTCTCT
CAGACCCCTCTCTGCGAGTATGCTAAACCTCCCTCTCTTTCTTATCCGCTGTCCCAT
GGCCAGCCTGGATGAATCTATCAATAAAACAACTAGAGAATGGTGGTCACTGAGACACTAT
AGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGGATCGGGTGTTACAGGTACAAGTAGGTA
TGTTGCAGAGGAAATAAATATCAAACTGTATACTAAAATTAATAA

FIGURE 162

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71180
><subunit 1 of 1, 205 aa, 1 stop
><MW: 21521, pI: 7.07, NX(S/T): 1
MLGAKPHWLPGLHSPGLPLVLVLLALGAGWAQEGSEPVLLGECELVVCEPGRAAAGGPGGA
ALGEAPPGRVAFAAVRSHHHEPAGETGNGTSGAIYFDQVLVNEGGGFDRASGSFVAPVRGVY
SFRFHVVVKYNRQTVQVSLMLNTWPVISAFANDPDVTREAAATSSVLLPLDPGDRVSLRLRRG
NLLGGWKYSSFSGFLIFPL
```

Signal peptide:

amino acids 1-32



FIGURE 163

GCTGTTTCTCTCGCGCCACCACTGGCCGCCGGCCGCAGCTCCAGGTGTCTAGCCGCCAGC
 CTCGACGCCGTCCCGGGACCCCTGTGCTCTGCGCGAAGCCCTGGCCCCGGGGGCCGGGCAT
 GGGCCAGGGGCGCGGGGTGAAGCGGCTTCCGCGGGGCCGTGACTGGGCGGGCTTCAGGC**AT**
GAAAGACCCTCATAGCCGCTACTCCGGGTCTGCGCGCGAGCGTCAGGCCGAGGCTGACC
 GGAGCCAGCGCTCTCAGGAGGACCTGCGCTGTGCGCGGAGGGGTCTGGGAGATGGGGCACT
 GGATCCAGCATCCTCTCCGCCCTCAGGACCTCTTCTGTGCTACCTGGCTCAATAGGTCCAA
 GGTGGAAGCAGCTACAGGTCATCTCAGTGCTCCAGTGGGTCTGTCTTCCCTTGTAAGTGG
 GAGTGGCCTGCAGTGCCATCCTCATGTACATATTCTGCACGTATTGCTGGCTCATCGCTGTG
 CTCTACTTCACTTGGCTGGTGTGTTGACTGGAACACACCCCAAGAAAGGTGGCAGGAGGTACA
 GTGGGTCCGAAACTGGGCTGTGTGGCGCTACTTTCGAGACTACTTTCCTCCATCCAGCTGGTGA
 AGACACACAACCTGTGACCAACCAGGAACATATCTTTGGATACCACCCCATGGTATCATG
 GGCCTGGGTGCCTTCTGCAACTTCAGCACAGAGGCCACAGAAGTGAGCAAGAAGTTCACAGG
 CATACGGCCTTACCTGGCTACACTGGCAGGCAACTTCCGAATGCCTGTGTTGAGGGAGTACC
 TGATGTCTGGAGGTATCTGCCCTGTGAGCCGGGACACCATAGACTATTTGCTTTCAAAGAAT
 GGGAGTGGCAATGCTATCATCATCGTGGTCGGGGGTGCGGCTGAGTCTCTGAGCTCCATGCC
 TGGCAAGAATGCAGTCACCTGCGGAACCGCAAGGGCTTTGTGAAACTGGCCTGCGTCATG
 GAGCTGACCTGGTTCCCATCTACTCCTTTGGAGAGAATGAAGTGTAAGAAGCAGGTGATCTTC
 GAGGAGGGCTCCTGGGGCCGATGGGTCCAGAAGAAGTCCAGAAATACATTGGTTTCGCCCC
 ATGCATCTTCCATGGTCGAGGCCCTCTTCTCCTCCGACACCTGGGGGCTGGTGCCCTACTCCA
 AGCCCATCACCCTGTTGTGGGAGAGCCCATCACCATCCCAAGCTGGAGACCCAACCCAG
 CAAGACATCGACCTGTACCACACCATGTACATGGAGGCCCTGGTGAAGCTCTTCGACAAGCA
 CAAGACCAAGTTCCGGCCTCCCGGAGACTGAGGTCCTGGAGGTGAAG**TGAG**CGCCAGCCTTCGGG
 GCCAATTCCTTGAGGAACCAAGCTGCAAACTACTTTTGTCTGTAAATTTGGAAGTGTC
 TGGGTGTCTGTGGGTATTTAAAAAGAAATTATAACAATTTTGTCAAACCAAAAAAAAAAAAA
 AA
 AAAAAAAAAAAAAAAAAAAAA

FIGURE 164

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184

><subunit 1 of 1, 388 aa, 1 stop

><MW: 43831, pI: 9.64, NX(S/T): 3

MKTLIAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRS
KVEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLI AVL YFTWLVFDWNTPKKGGRRS
QWVRNWA VWR YFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAF CNFSTEATEVSKKFP
GIRPYLATLAGNFRMPVLREYLMSSGGICPVSRDTIDYLLSKNGSGNAIIIVVGAAESLSSM
PGKNAVTLRNKGFVKLALRHGADLVPIYSFGENEVYKQVIFEEGSWGRWVQKKFQKYIGFA
PCIFHGRGLFSSDTWGLVPYSKPITTVVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDK

HKTKFGLPETEVLEVN

Important features of the protein:

Transmembrane domain:

amino acids 76-97

N-glycosylation sites.

amino acids 60-63, 173-176, 228-231

N-myristoylation sites.

amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-245, 318-323, 378-383

[illegible]

FIGURE 166

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213
><subunit 1 of 1, 368 aa, 1 stop
><MW: 42550, pI: 9.11, NX(S/T): 1
MGLLAFLKTQFVLHLLVGFFVVSGLVIN FVQLCTLALWPFVSKQLYRR LNCRLAYSLWSQLV
MLEWWSCTECTLFTDQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGLGSSKVLAKK
ELLYVPLIGWTWYFLEIVFCRRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHR
VSMEVAAAKGLPVLKYHLLPRTKGF TTA VKCLRGTVAAVYDVTNLNFRGNKNPSLLGILYGKK
YEADMCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLLN
FLSWATILLSP LFSFVLGVFASGSPLLIL TFLGFVGAASF GVRRLIGESLEPGRWRLQ
```

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 307-323, 335-352

Tyrosine kinase phosphorylation sites.

amino acids 160-168, 161-169

GATATTCCTTTATTTTAAAGAATCTGAAGTACTATGCATCACTCCCTCCAATGCTCTGGGGCA
GCCACCAGGCATATTCATCTTTGTGTGTGTCTTTCTTTGCTTTAGCACTGGGGCACTTCTT
GCTTATTTCTTTGTGTAGAAAGGGGCTCAGTTTGCTTGTGGGGTTGGTGGCAGGCAGGCCG
GCTTACGCCTGATACGCCCTGGGTTAGAAAGGAAGGAAGATAAACTTTTATACAAATGGG
GATAGCTGGGGTCTGAGACCTGCTTCCTCAGTAAATTCCTGGGATCTGCCTATACCTTCTT
TTCTCTAACCTGGCATACCCTGCTTAAAGCCTCTCAGGGCTTCTCTCTGTTCTTAGGATCAA
AGTATTTAGAGCTACAAGAGCCCTCATGGTCTGGCCCCCTGCCCCCCCTGGCCAGCTTCATTGT
ACATGTGGTGTCTCTTGTGTTCTCTGTAAATGTGGTATGCCATGGGGTCTTTGCACAAGCCT
TTCCTCTTTGGCTGGACACTGTTTCCTCGCCCCCCCCATACTCTTCTTACTTAATATGTAGTC
ATCCCTGCAGATTCAATCTTAACATCATTTTCTCCAGGATCTTGGCTGACAGAATCTCAT
CTTGTTTAATGCTCTCATAGACCACTGTTTTCCTTTTGGCAGCACTTGCCTACTCAGTTGTA
TCTTTATGTGCGTTTGTGGTTGTATGGTGTGTCTGTTCCCGAGAATGCCAGCTCTGAGC
TGCGTGAGGGTCAAGGGCATTGCTGTGCCTGCCAGGTATAGTGCCTACATGTGGTGGGTGCT
CATGTTTTAGAGACTAAATGGAGGAGGAGATGAGGAAAAGATTGAAATCTCTCAGTTCACCA
GATGGTGTAGGGCCAGCATTGTAATTCACACGTTGACTGTGCTTGTGAATTATCTGGGGA
TGCAGGTCTTGATTAGTAGGCCAGGTTGGGCATCTCTAACAACTCCCAGTGATGCTGA
TGCTGGTCTATGAACATATACTAAATAGTAAGAATCTATGGAGCCAGGCTGGGCATGGTGGC
TCACACCTATGATCCCAGCACTTTGGGAGGCTGAGGCAGGCTGATCACCTGGAGTCAGGATT
TCAAGACTAGCCTGGCCAACATGGTGAACCCCATCTGTACTAAAAATACACAAATTAGCTG
GGCATGGTGGCAGCTGCTGTAGTCCCAGTACTTGGGAGGCTGAAGCAAGAGAATCGCTTG
AACCTGGGAGCGGAGGTTGTAGTCAGGCCGAGATCAGGCCACTGTATTCCAACCGGGTGAC
AGAGTGAGACTCTATGTCCAAAAAATAAAAAA

FIGURE 168

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234
><subunit 1 of 1, 143 aa, 1 stop
><MW: 15624, pI: 9.58, NX(S/T): 0
MHHS LQCPGAATRH IHL CVCF SFALALGH FLLISLVGKGLSLSCGVGGRQAGLR LIRPWVRR
EGKINFY TNGDSWGLR PASSVKFLGSAYTFFSLTWHTLLKASQGFSLFLGSKYLELQEPSWS
GPCPPGQLHCTCGVLLSFL
```

Important features of the protein:**Signal peptide:**

amino acids 1-28

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 169

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGA
 TTATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAAGCCAC
 AGGAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTAC
 ATTAAAATCTGTTTTTTTGTCTCTTGTAAGTACGCTTTTACCTTCCTAACACAGAGGATCTGT
 CACTGTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCAC
 ACCGTCCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTC
 ACCAACTGTCTCAGCTCTGGAGGCACTGACTCGGGCAGTGCAAGGTAGCTGAGCCTCTTGGTA
 GCTGCGGCTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTGACAAGGCCGAAGATT
 CATAGCGGATGGCTCCCACTGCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCCTGGGG
 CCAGGACGGGCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCGCCCAT
 CTAACCTTTTCATGTCTGTCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGG
 AACCAGAGCTTGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTT
 TATGTGACAGGACTTGCACTTCTCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCA
 GGGAAAGAACTTGTGCCAAATTATGGGTCAGAAAAGATGGAGGTGTTGGGTTATCACAAGGC
 ATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACT
 CGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCGATCCACGTACCAGCTGCTG
 AAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAG
 CCAGGGGCAGCCGCTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTC
 TGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATG
 GCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCG
 AACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCT
 CAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGT
 TCAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAA
 TCCAAACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATG
 CCAACATTTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTG
 GCCAACATGGTGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGT
 GTGTGCTGTATCCAGCTACTCGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGT
 GAAGGAGCTGAGACAGGAGAATCACTTCACTGAGCAACACAGCGAGACTCTGTCTCAGA
 AAAAAATAAAAAAGAATTATGGTTATTGTAA

FIGURE 170

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71277
><subunit 1 of 1, 109 aa, 1 stop
><MW: 11822, pI: 8.63, NX(S/T): 0
MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTA
SPCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD
```

Signal peptide:

amino acids 1-15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

CGCGGCCCGCGAGTCCGAGACCTGTCCCAGGAGCTCCAGCTCACGTGACCTGTCACTGCCTC
C CGCGGCCCTCCTGCCCGGCC**CATG**ACCCAGCCGGTGCCCGGCTCTCCGTGCCCGCCGCGCT
GGCCCTGGGCTCAGCCGCACTGGGCGCCGCTTCGCCACTGGCTCTTCTGGGGAGGCGGT
GCCCCCCATGGCGAGGCGCGGAGAGCAGTGCCTGCTTCCCCCGAGGACAGCCGCTGTGG
CAGTATCTTCTGAGCGCTCCATGCGGGAGACCCGGCGCTGCGAAGCCTGAGGCTGCTGAC
CCTGGAGCAGCCGAGGGGGATTCTATGATGACCTGCGAGCAGGCCAGCTCTTGGCCAACC
TGGCGCGGCTCATCCAGGCCAAGAAGGCGCTGGACCTGGGCACCTTACGGGCTACTCCGCC
CTGGCCCTGGCCCTGGCGCTGCCCGCGGACGGGCGCGTGGTGACCTGCGAGGTGGACGCGCA
GCCCCCGAGCTGGGACGGCCCCCTGTGGAGGCAGGCCGAGGCGGAGCACAAGATCGACCTCC
GGCTGAAGCCCGCCTTGGAGACCTGGACGAGCTGCTGGCGGCGGGCGAGGCCGGCACCTTC
GACGTGGCCGTGGTGGATGCGGACAAGGAGAAGTGTCTCCGCTACTACGAGCGCTGCCTGCA
GCTGCTGCGCAGCCGGAGGCATCTCGCCGCTCTCAGATCCTGTGGCGCGGAAGGTGCTGC
AACCTCCGAAGAGGGGAGCTGGCGGCGGAGTGTGTCGAACACTAAACGACGCTCCGGCGG
GACGTGAGGCTACATCAGCCTCTCGCCCTGGGCGATGAGTCACTTGTGCCCTCAAGAT
C**TAG**GGCTGGCCCTAGTGAGTGCGCTCGAGGAGGGTGTGCTGGGAACCCAGGAATTGAC
CCTGAGTTTTAAATTGAAAAATAAAGTGGGGCTGGGACACAAAAA

FIGURE 172

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71282
><subunit 1 of 1, 262 aa, 1 stop
><MW: 28809, pI: 8.80, NX(S/T): 1
MTQPVPRLSVPAALALGSAALGAATGFLGRRCPPWRGRREQCLLPEDSRLWQYLLSRS
MREHPALRSLRLLTLEQPQGDMMTCEQAQLLANLARLIQAKKALDLGTFTGYSALALALAL
PADGRVVTCEVDAQPPELGRPLWRQAEAEHKIDLRLKPALETLDELLAAGEAGTFDVAVVDA
DKENC SAYYERCLQLLRPGGILAVLRVLRGKVLQPPKGDVAAECVRNLNERIRRDVRVYIS
LLPLGDGLTLAFKI
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

Transmembrane domains:

amino acids 8-30, 109-130

N-glycosylation site.

amino acids 190-193

Tyrosine kinase phosphorylation site.

amino acids 238-246

N-myristoylation sites.

amino acids 22-27, 28-33, 110-115, 205-210, 255-260

Amidation sites.

amino acids 31-34, 39-42

[illegible]

FIGURE 174

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286

>subunit 1 of 1, 671 aa, 1 stop

>MW: 74317, pI: 7.61, NX(S/T): 0

MPHAFKPGDLVFAKMGYPHWPARIDDIADGAVKPPPNKYPIFFFGTHETAFLGPKDLFPYD
KCKDKYKGPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDEDGRG
VMAVTAVTATAASDRMESDSDSKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSE
EENSESSSESEKTSQDQFTPEKKA AVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPV
AMARSASSSSSSSSSDSDSVKKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEV
DRISEWKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGD
ELREDDEPVKKRGRKGRGRGPPSSSDSEPEAELEREAKKSAKPKQSSSTEPARKPGQKEKRV
RPEEKQQAQKPVKVERTRRKSEGFSDMRKVEKKKEPSVEEKLQKLHSEIKFALKVDSPDVKRC
LNALEELGTLQVTSQILQKNTDVVATLKKIRRYKANKDVMKAAAEVYTRLKSRVLGPKIEAV
QKVNKAGMEKEKAEEKLAGEELAGEEAPQEKAEDKPSTDLAPVNGEATSQKGESAEDEKEH
EGRDSEEGPRCGSSEDLHDSVREGPDLDPRPGSDRQERERARGDSEALDEES

Signal peptide:

amino acids 1-13

FIGURE 175

GTTGGTTCTCCTGGATCTTCACTTTACCAACTGCAGATCTTGGGACTCATCAGCCTCAATAATTATATTAATAATTA
 ACACCATTTTGAAGAGGACATTTGTTTTTCATC**ATGA**ATGCTAATAAAGATGAAAGACTTAAAGCCAGAGCGAAGA
 TTTTCACTTTTTTCTGCTTTTGATGATGCTAAGCATGACCATGTTGTTTCTTCCAGTCACTGGACCTTTGAAGCA
 AATATATCCAGAGCTCAGGCTAACCTACAAGAGCTTGTGCTTTCAAATAGCTGATTTCCCTTTTTTGGGTTTCATC
 AGAAGAGCTGGATTTTCAAATCTTTCTCTTAGATGAGGAAGAGGCGAGGCTGCTCTTTGGGAGCCAAAGACCATCT
 CTTTCTACTCAGTCTGGTTGACTTAAACAAAATTTTAAAGAAATTTATGGCCTGCTGCAAAAGGAAACGGGTGGA
 ATTTATGTAATTTAGCTGGGAAGATGCCAATACAGAATGTGCAAATTTCAATCAGAGTACTTCAGCCCTATAACAA
 AACTCACATATATGTGTGGGAATGGAGCATTTTCATCCATATGTGGGTATATGATCTTCAGGTCTACAAGGA
 GGATATTTATTTCAAATAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCCCTTCGATCTCAGCAGCC
 TTTTGCTTCAGTAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCCCTTGGCAAGATACTGCAAT
 CACTCGATCCCTTGGGCTACTCATGACCACCACTACATCAGAAGTACATTCAGAGCACTACTGGCTCAATGG
 AGCAAAATTTTATGGAATTTCTTCATACAGACACCTTACAATCCAGATGATGATAAAATATATTTCTCTTTCTG
 TGAATCATCTCAAGAAGGCGAGTACCTCCGATAAAACCATCTTTCTCGAGTTGGAAGAGTTTGAAGAAATGATGT
 AGGAGGACAACGAGCCTGATAAACAAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTCCTGGGAAG
 TGATGGGGCAGATACTTACTTTTGTAGCTTCAAGATATTTATTTACTCCCCACAAGAGATGAAAGAAATCCTGT
 AGTATATGGAGTCTTTACTACAACAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTATAGCATGGCTGACAT
 CAGAGCAGTTTTTAAATGGTCCATATGCTCATAGGAAGTGCAGACCATCGTGGGTGCAGTATGATGGGGAAT
 TCCTTATCCAGGGCTCTGATACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCAGATTTTCTCAGA
 TGTATGTCATCAGTTTCTGTAAGAGCCGACTCTGTGATGTATAAGTCCGATACCCAGTTGCAGAGGAGCAACCGTT
 CAGAGAAGATCAATGTGGATTCAGAGCTGACACAGATAGTGTGGATCATGTTCATTGCAGAAGATGGCCAGTACGA
 TGTAATGTTTCTTGGAAAGACATTTGGAATCTGTCCCTCAAACCTGTGCAGCATTTCAAGAGAAAGTGGAAATGTCT
 AGAGGTAGTGTGAGGAGGTTGCAGATATTCAGCACTCATCAATCATCTTGACATGGAATTTGTCTGGAAGCA
 GCAACAAATGTACATTTGGTTCCCGAGATGGAATAGTTGACGCTCTCTGCAACAGATGCGACACTTATGGGAAAGC
 TTAGCAGAGCTTGTCTTGTCCGACAGACCCCTACTGTGCTGGGATGGAAATGCATGCTTCGATATGCTCTCTAC
 TTTCAAAGAGAGCTAGACCGCAGATGTAATAATGAGCCCAATCACCAGTGTCTGGGACATCGAAGACAG
 CATTAGTCTGAATCTGCTGATGAAAAGGTGATTTTGGCATTGAAATTTAACTCAACCTTTCTGGAATGTATTACC
 TAATCCCAACAGCACTATTAATGGTATATCCAGAGGTACGGGATGAGCATCGAGAGGAGTTGAAGCCCGA
 TGAAGAAATCATCAAAACGGAATATGGGCTACTGATTGGAAGTTGAGAGAAGAGGATTTCTGGATGTATTACTG
 CAAAGCCAGAGGACACATTTCTCCACACCATAGTGAAGCTGACTTTGAAATGTCAATGAGAATGAACAGATGGA
 AATACCCAGAGGACAGCATGAGGAGGGGCGAGGTCAAGGATCTATTGGCTGAGTCAAGGTTGAGATACAAAGA
 CTACATCCAAATCTTTAGCAGCCCAAACTTCAGCCTCGACCACTACTGCGAACAGATGTGGCAAGGAGAGGCG
 GAGACAGAGAACAAGGGGGGCCAAAGTGAAGACACATGCAGGAATGAAGAAAGAACGAAATCGAAGACATCA
 CAGAGACCTGGATGAGCTCCCTAGAGCTGTAGCCACG**TA**GTCTTCTACTTAATTTAAAGAAAGAAATTCCTTACC
 TATAAAAAATTCCTTTCTGTTTGTATATCCCTTATAGATTAATCATAAATGCTTCCCATGGATTTTCTGAAGG
 CACAAGACAAATATCTGAATAAGCAATATGTGATGAATATAAGAAAGGGCAAAATTCATTGTAACAGATTTT
 CCAAGAACAAATCTTGCACAAGCAAGTATAAGAAATATCCTAAAAATAGGGGGTTTACAGTTGTAATGTGTTTA
 TGTTTTGAGTTTGGAAATTTATGTCAATGAATAGTTGAGCTAAGCAAGCCCGAAATTTGATAGTGATATAGGT
 GCTTTATTCCTCGAATGTCCATTAAGCATGGAATTTACCATCAGTTGTGCTATGTTCTTATGAACAGATATAT
 CATTCTATTGAGAACAGCACTACCTTTGGTAGGGAATAAGAGGTGAGACACAAATTAAGACAACTCCCATATC
 AATCGAATTTCTCAGTGAGCCATTCACTCTGGAGAATGGTATAGGAATTTGAGAGGTTGCAATATTTCTTTCT
 TGCCCATCTGGGTTTAAATTTAGTGTACTACAACATTTGATTTACTGAAGGGCACTAATGTTTCCCGCAGGATTTCT
 ATGTAAGTACAGGAGTAACAGCTTTCACAGAGAGAAGTTGGTGCTTATGTTTATGTTATGATATATACTAA
 GCTCTACAGGACAGAAATCTTATAAAATACTTTAATAAGATATGGGAAATATTTTATAAAGAACAGGAAACA
 TAATGATGATAATGATCATTGATGGGAGGCACTGAGATGGGATTTGTAAGAAGAGCAAGGAGCAAGCCAT
 AAATCTGCTGCTTTGGGGAAGCACTCATCCCCATGAAAGGAAGAACAACTCAAAATAAAGTGAGAGTATGTA
 TGGAGCTCTTTTCTAGTGGTATAAGTAGCTGCCAATTTGTAATCTACTGTTTAAAAAAATCTAGATTTATAACA
 AACTGCTAGCAAAATCTGAGGAACACAATAATCTTCTGAAGAATCATAGGAAGAGTAGCATTTTATATAACC
 AATGATTTTTCAGTATATATTTCTCTCTTTTAAAAAATATTTATCATCTCTGTATATTTCTTTTACTCGC
 GTTTATCTCTCTGATATATTTGGATTTTGTATATATTTGAGTGAATAGGAGAAACATATATTAACACACGA
 CAATTAAGAAATTTGATTTTGGGGAGTGGGATATATATTTGTAACACAGACAGTGTAAATTTTAAAC
 AACGGAAGGGTTAAATTAACCTTTTGACATCTTCACTCAACCTTTTCTCATTTGCTGAGTTGCTGTTGTAAT
 GTAGTATGTTTTGTAATTTAAACATAAATAGCCCTGCTACATGT

FIGURE 176

```

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883
><subunit 1 of 1, 777 aa, 1 stop
><MW: 89651, pI: 7.97, NX(S/T): 3
MNANKDERLKARSQDFHLFPALMMLSMTMLFLPVTGTLKQNIPLRLKLT YKDLLLSNSCIPFL
GSSEGLDFQTLLLLDEERGRLLLGAKDHI FLLSLVDLNNKFNFKIYWPAAKERVLCCKLAGKDA
NTECANFIRVLQFPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFD
PQQPFASVMTDEYLYSGTASDFLGKDTAFTRSLGPTHDHHYIRTDISEHYWLN GAKFIGTFF
IPDTYNPDDDKIYFFFRESSQEGSTSDKTI LSRVGRVCKNDVGGQRS LINKWTTFLKARLIC
SIPGSDGADTYFDELQDIYLLPTRDERNPV VYGVTFTTSSIFKGSAVCVYSMADIRAVFNGP
YAHKESADHRWVQYDGRIPYPRPGTCPSK TYDPLIKSTRDFPDDVISFIKRHSVMYKSVYPV
AGGPTFKRINVDYRLTQIVVDHVIAEDGQY DVMFLGTDIGTVLKVVVISKEKWNMEEVVLEE
LQIFKHSSIIILNMELSLKQQQLYIGSRDGL VQLSLHRCDTYGCACADCLARDPYCAWDGNA
CSRYAPTSKRRARRQDVKYGDPIQCDWIEDS ISHETADEKVI FIEFNSTFLECIPKSQQA
TIKWYIQRSGDEHREELKPDERIIKTEYGL LIRSLQKKDSGMYCYCAQEHTFIHTIVKLTLN
VIENEQMENTQRAHEEGQVKDLLAESRLRY KDYIQILSSPNFSLDQYCEQMWHREKRQRN
KGGPKWKHMQEMKKRNRHRRDLDELPRAVAT

```

Important features of the protein:

Signal peptide:

amino acids 1-36

N-glycosylation sites.

amino acids 139-142, 607-610, 724-727

Tyrosine kinase phosphorylation site.

amino acids 571-576

Gram-positive cocci surface proteins 'anchoring' hexapeptide.

amino acids 32-37

FIGURE 177

CCCTGACCTCCCTGAGCCACACTGAGCTGGAAAGCCGAGAGGTCATCCTGGAGCATGCCACCACGGGGGAGCAGA
 CAACCTCCACAGTGAAGCTGGGAGCAGAGCTGAAGCTGTTCTTCAGGAGCTGGGTGATATTTTCCCCACCCAC
 CTGACAGTTTTCAGCCAGCAGGGACTGATCAGGTGTGTCTCGAGTGGGGAGAGAGGCGCTGGCCGCAAGA
 GTGGCTGGAGAAAGAGGTTTCAGCGCTTGACAGCCGAGCTGCCCGTGAATACAGATCCAGAACCATGGGCATC
 GTGTGAGGTGGGGGGGACAGGTGTCATGTGCACTTCTTGCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG
 AGCCATTGAGGGTGTATGAGACTACAGAGGGGAGGGAAAGGTATTTAAAGGTAAACAGTGTGGCACAATAGTTAA
 GAGCAGAGTTTTTGAGCTGAGCCGACATAGGTTCAAAATCTCTCTGTTGCTTCTCTGTTCTGATGCCCCAGT
 AAGGAGTGACTTAACCTCTCTGGACTTCAATTTCTCATCTAAAGTAGGGCCAAATAATAGCACCCACCTCAT
 AGGGAAGATTAATGACATAATGTATGTATGATGCACTCACTAGCAAAGTACCAGTCCCATAGTAAGTCTGCCCCAG
 TATTTCCACCCACCCCTGTTCTCTGCTTCCCAACAGGTACTGCAACGACTGGAGCAGAGGCGGCAGCAGGCTT
 CAGAGCGGGAGGCTCCAAGCATAGAACAGAGGTTACAGGAAGTGCGAGAGAGCATCCGCGGGGCAAGGTGAGCC
 AGGTGAAGGGGGCTGCCCGCTGGCCCTGCTGCAGGGGGCTGGCTTAGATGTGGAGCGCTGGCTGAAGCCAGCCA
 TGACCCAGGCCAGGATGAGGTGGAGCAGGAGCGGCGCTCAGTGAGGCTCGCTTCCAGAGGGACCTCTCTC
 CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGGAATGTGAGGAGAGCGGAGAGCTCTTTGAGGAGCTGCC
 CCCAAGCCCTGGCCACGAGGGCCCTCCCTGCCCTGCACACGTGGTATTTTCGCTATCAGGCGAGGGCTGAGGAT
 AGCTGACAAATCAGGAGGGTGAATGGCTGAGGCTCATAGAGAGGGAGATCTGACGAATGGGTCAAGCTCGGA
 ACCAGCAGCGGAGGTAGGCTTTGTCCTGAGCGATATCTCAACTTCCCGGACCTTCCCTCCAGAGAGCAGCC
 AAGACAGTGACAATCCTCGCGGGCAGAGCCACAGCATTCCTGGCAACAGCCCTGTACAGCTACACCAGGACAGA
 GTGCAGAGAGCTGAGCTTCCCTGAGGGGCACTCGCTGCTGCCCGGGCCAGCATAGTGAATGATGAGCG
 GTTTCGAGGGGAGAAATTTGGGGCCGCTGTGGGGCTTCCCTCCCTGCTGTGGTGAAGAGCTGCTTGGCCCGG
 CAGGCGCCAGCTGAACCTTCTGAAGCTGAACAGATGCTGCCGTCCCTCTCTCTCCAGCTTCTCTCCCACTGCAC
 CTACCTCTGTGTGATGGGCCCTGCACCTGTCTGCTGGGCAAGGAGCCCTGAGCTTCTCCCTGGTCTCTCTGG
 ACATGATGCCACCTGCACCTCAGGCCGATGCTGCACCACTCCCGCGGGCTAAGAGCCCGGATCTGGCCACC
 CAGATCCCTCACCTGAGAGGCCAGGGAGGCTTGCACCCAGTATGCTGCTGCTGCTTCCCTATCTTAAGCTGTCA
 GACCCATCAATGATCCAGAGCAGACAGCCAAAGCTGGAATCGCCCTTATTTCCACCTCAGCTCCAGAGGT
 GGAACCTTGGCCCTTCCCTATTTCTAGAGCTGGAAACCCACTCTCTTTTCCATTTCTCATCTCTAGGACC
 GGAATCTACTCTCTCTCTGATGACCTTATCTAGGTTGGTGAATGCTGAATCTCTGGGGCTGGAAACC
 ATCCATCAAGTCTCTAGTAGTCTGGCCCACTCTTCCCAACCTGGCTCCATGACCCACCCACTCTGGATG
 CCAGGTCAGTGGGGTTGGGCTGGGAGAGGAAAGGCTTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG
 CTGTAATGGCTGAGCGGATTTATTGACAAATGAATAAGGGCACGAAGGCCAGGCCAGGGCTGGGCTCTTGTG
 CTAAGAGGCGAGGGGCTACGGTGTATTTGCTTTAGGGGCCACACCGGCCAGGGGCTGCTCCAGCTGCCAC
 GCTCTATCATATGGAGCAGGTTGTTGGGAAGGCCGGGCGAGCAGCTGTTGAGGCGAGGGGAGGAGAAGAGAC
 TGAGGGGCTGTGACCTCTCCTGAGGCCCCAGCCTGAGACTGTGCACTCCAGGTTGAAGTAGAGCTGGTCCCT
 AGCTGGGGGCGAGTGTCTGACCTGAGGGGAGGGCTTTCAAGCCCAACCAAGCCCTGGCCCTGCGAGCTGTAG
 TCCATCAGCACAATGAAGGAGCTTGAGAAAGAGGAAGAATAACACTGTTGCTTCTGTCTCAAGCTGTGTCCAGC
 TTTTCCCTGGGGCTCCAGGACTTCCCTACCTCCACCAACCAAGGATTTATAGCAAGGCTTAGCCTGC
 AGTTTACTCTGGGGTTTCAAGGAGCCGAAGGCTTAATAGTTTAAAGTAGGTGATGGGAAGATGAGATTACCTCA
 TTTAGGGCTCAGGCACTCACTCACACTCTCTGCTGCTGCTGTTGGTAGAGCAATGAAGAGCCAGCCAGTGC
 TCAACAATGAGAGCAGGATAGGTCCTATCAGTGCCGCCAGAGTAGAGAGCAATAAGAGCCAGCCAGTGG
 AGTCCCGGCTGTGTTTCTACCTGTTGATCAGAAGTCTGTTGCTTGGCTGCCATTTGCTCTTGAAGTGG
 GCAGCCCTGGGCTGGGCCCTCCCTCCGGCCCTCAGTGTGGCTGTGCAAGAGCTCTGGGGTTCCCTTCAAGTG
 CAGCAGGGTTAGGCTGTCTGCTGAGTCTCCTCAGTCTGATGAGGCTGGCTAGGACCTGGGGCTGTGGCC
 TCTCAGGGGCGACCTCTCATGGCAGGCATCCCTGCTTGGGCTGCCCTCCCCAGAGCCCTGACACCCCTGCT
 GTGCTGTCCCAAGGAGCCAGCTCCTGCTGTGGGGAGGCATCAGGCTGTGTCGATGATAGGCTGCTGCT
 TCTCAATGTGTGACCCGGGACCTGGGAGGGGAGGAACACTGGGTTTAGGACCAACTCAGAGGCTGCTGT
 GCCTCCCTCTGACCAAGGACCTCTGAGTTTGGTGGCTACTTCCCTTGGGCTTAAGTAGGGGAGGCTTCTC
 AGATTGTGGGCACTTGTGTAGCTGACTTCTGCTGGAGCTCCAGTCCAGGAGGAAGAGCCAGGCCAATTT
 TGGGATCAGGTGCTGATCACTGGCCCTTACCTCAGCCCTTCTCCCTGAGACCTGCCCACTGCCCA
 CAGAGAACACAGTGGTCTCCCTCTCCGGGGCGGCTTTTCTCTCTTGGAGCTCCCTGAGCAGCAAGTGAAG
 GTCTTCTGCTGGGCTGCAATGATGACGAAGGGCTGCAGAGCCAGTGCATGTGATGTGGGAGGGGCTC
 CGTCTGCGAGGCTGGAGGTGGCATCCCACTGCACAGCAGGAGGAGGGAGTGAGGGTAACATTTCCATTTCCCT
 TCTATGTTTGTGTTTCACTGCTTCTCAGCATGCTCTTAAACCCAGAGGCCCAATTTCCCGAGGCCCATTT
 TTTCTGTCTTATCTAATAAATCAATATTAAG

1001571.1.11111

FIGURE 178

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401
><subunit 1 of 1, 370 aa, 1 stop
><MW: 40685, pI: 4.53, NX(S/T): 0
MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLEQRRQQASEREAPSIEQRLQEVRESIRRA
QVSQVKGAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELS
DFEECEETGELFEEPAPQALATRALPCPAHVVFQYQAGREDELTITEGEWLEVIEEGDADEW
VKARNQHGEVGFVPERYLNFPDLSLPSSQSDSNPCGAEPATAFLAQALYSYTGQSAEELSFP
EGALIRLLPRAQDGVDDGFWRGEFGGRVGVFFPSLLVEELLGPPGPPELSDPEQMLPSPSPPS
FSPAPATSVLDGPPAPVLPDGKALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT
```

180/249

FIGURE 179A

CACAGGGAGACCCACAGACACATATGACACGAGAGAGACAGAGGAGGAAAGAGACAGAGACAAAGGCCACAGCGGAA
 GAAGGCAGAGACAGGGCCAGGCACAGAGAGCGGCCACAGACAGAGTCTACAGAGGGAGAGGCCACAGAAAGCTGCAGA
 AGACACAGGCAGGAGAGACAAAGATCCAGGAAAGGAGGGCTCAGGAGGAGAGTTTGGAGAAGCCAGAGCCCTGG
 GCACCTCTCCACAGCCCAAGGACTAAGTTTTCTCCATTCTCTTAAACGGTCTTCAGCCCTCTCGAAACTTTGGC
 CTCGATTTGGCAGAGTCCAAAGCCCCAGGCTACAGAGAGAGAGCTTCCAAAGCTAGGGTGTGGAGGACTTTGGT
 GCCTTAGACGGCCTCAGTCCCTCCAGCTGCAGTACCAGTGCCATGTCCTCCAGACAGGCTTCGATCCCGGGAGGGG
 CTGTGGACGGGCTGGCTGTGGGGAGCCAAACCTGCCTCCTGCTCCCATTTGTGCCGTCTCTGGCTGTGTGTG
 CTGTGTTCTGTCTACTGCTGGCTCTCTCTGCCCCAGCCCGGCTGGCCAGCCCTCTCCCCGGGAGAGAGAT
 GCTGTTTTCCAGAGAAGCTCAACGGCAGCGCTCTGCTGGCTGGGGCGCCCTGCGAGGCTTTGTGCCGTGTGA
 GGCCTTTGGGGAGACGCTGCTACTAGAGCTGGAGCAGGACTCCGGTGTGCAAGGTTCAGAGGGGTGACAGTGCAGTA
 CCTGGGCCAGGCGCTGAGCTGCTGGGTGGAGCAGAGCCTGGCACCTACCTGACTGGCACCATCAATGGAGATCC
 GGAGTCGGTGGCATCTCTGCACTGGGATGGGGAGCCCTGTTAGGCGTGTACAATATCGGGGGGTGAACATCCA
 CCTCCAGCCCCGTGGAGGAGGCCACCCCTAACTCTGCTGGGGAGCTGGGGCTCACATCCTACGCCGAAGAGTCC
 TGGCCAGCGGTCAAGGTCCCATGTGCAACGTCAAGGCTCCTCTTGGAAAGCCCCAGCCCCAGACCCGAAGAGCCAA
 GCGCTTTGCTTCACTGAGTAGATTTGTGGAGACACTGCTGCTGGCAGATGACAAGATGGCCGCTTCCACGGTGC
 GGGGCTAAAGCGCTACCTGCTAACAGTATGGCAGCAGCAGCCAGGCGCTTCAAGCACCCAAAGCATCCGCAATCC
 TGTCACTTTGGTGGTGAATCGGCTAGTGATCTGGGGTCAAGCCAGGAGAGGGCCCCAAGTGGGGCCCACTGCTGC
 CCAGACCTCGGCAGCTTCTGTGCTGGCAGCGGGGCTCAACACCCCTGAGGACTCGGGCCCTGACCACTTTGA
 CACAGAGCTTTGTTTACCGCTCAGGACCTGTGTGGAGTCTCCACTTGGCAGACCTGGGATAGGCTGATGTGGG
 CACCGTCTGTGACCCGGCTCGGAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAAGGCTGATGCTGTCTA
 TGAACCTGGGTGATGCTTCAACATGCTCCATGACATGCTCCAGGCACTGCATCACTTTGAAATGGGCTTTGTGAGAC
 CTCTGCCATCTGATGGCCCTGTGATGGCTCATGTGGATCTCTGAGGAGCCCTGGTGTCCGCTCAGTGGCCGCTT
 CATCACTGACTTCTGGCAACCTGTATGGGCACTGTCTCTAGACAAACAGAGGCTCTCAATTTGATCTGCGCTGT
 GACTTTCCTGGCAAGSACTATGATGCTGACCGCCAGTGCAGGCTGACCTTTCGGGCGGCATCAGCCCATTTGCC
 ACAGTCTGGCCCGCCCTGTGCTGCCCTCTGGTGCTCTGGCCACCTCAATGGGCACTGCGATGCTGCCAGACAAACA
 CTGGCCCTGGGCGCACTGGACACCTCTGGGCGCCGCACAGGCTGATGGGTGGTGGTCTGCTCCACATGGACCA
 GCTCCAGGACTTCAATATTCACAGGCTGGTGGCTGGGGTCTTGGGAGCCATGGGGTGAATGCTCTCGGACCTG
 TGGGGTGGTGTGCTGCTTCCGAGACTGCAACGAGGCGCTTCCCCGGAAATGGTGGCAAGTACTGTGAGGG
 CGCGCTACCCGCTTCCGCTCCTGCAACACTGAGGACTGCCCACTGGCTCAGCCCTGACCTTCCGCGAGGAGGA
 GTGTGCTGCTACACCCACCGCACCAGCTCTTCAAGAGCTTCCAGGGCCCATGAGTCTGGGTTCTCGCTACAC
 AGGCTGGCCCCCAGGACCAAGTCAAACTCACTGCCAGGCCCGGCGCACTGGGCTACTACTATTTGCTGGAGCC
 ACGGGTGGTAGATGGGACCCCTGTTCCCCGGACAGCTCCTCGGTCTGTGTCAGGCGCCATGCATCCATCTGG
 CTGTGATCGCATCATTTGGCTCCAAGAAGAAGTTTGACAAGTGCATGGTGTGCGGAGGGGACGGTCTGTGGTGCAG
 CAAAGCAGTCAGGCTCCTTCAGGAATTTAGGTACGGATAACAACATGTGGTCACTATCCCCGGGGGCCACCA
 CATTTCTGTCCGCGAGCAGGAAACCTTGGCCACCGGAGCATCTACTTGGCCCTGAAGCTGCCAGATGGCTCTTA
 TGCCCTCAATTGTTGAATACACGCTGATGCCCTCCCCACAGATGTGGTACTGCTGGGGCAGTCAGCTTGGCTA
 CAGCGGGGCACTGCAGCCTCAGAGACACTGTCAAGCCATGGGCCACTTGGCCAGCCTTTGACACTGCAAGTCTC
 AGTGGCTGGCAACCCCCAGGACACAGCCTCCGATACAGCTTCTTGTGCGCCGCGGACCCCTTCAACGGCCAGC
 CCCCCTCCACAGGACTGGCTGCACCAAGAGACAGATTTCTGGAGATCCTTCCGCGGCGCCCTGGGCGGGCAG
 GAAATTAACCTCACTATCCCGGCTGCCCTTTCTGGGACCGGGGCTCGGACTTACGTGGGAGAAGAGAGAGCTT
 CTGTTGCTGCTCCTCATGCTAAGACTCAGTGGGGAGGGGCTGTGGCGTGGAGACTGCCCTCCTCTCTGCCCTAAT
 GCGCAGGCTGGCCCTGCCCTGGTTTCTGCCCTGGGAGGCACTGATGGGTTAGTGGATGGAAGGGGCTGACAGAT
 AGCCCTCATCTAACTGCCCTTCTGCCCCGTGCGGGTCAAGAGAGGAGGGGGAAGGAGGAGGCGCTGGGGCC
 CAGTGTGTTATTTATTTAGTATTTATCTACTTTTATTTAGCACAGGAGAGGGGACAGGACTAGGGTCTCGGGGAA
 CCTGACCCCTGACCCCTCATAGGCTCACCCTGGGGCTAGGAAATTCAGGGTGGTGGTGTATAGGTATTAAGTGGTG
 GTGTGTGTGTGTGTGTGTGTGTGTGAAATGTGTGTGTGTATGTATGAGTCAACCTGTTCTGTCTTCTCTC
 TTCTCGAATTTTATTTTGGGAAAGAAAGTCAAGGGTAGGGTGGGCTCTCGAGGATGAGGGATATCTPTT
 TTTTTTTTCTTCTTCTTCTTCTTCTTCTTCTTGTGAGACAGAATCTCGCTCTGCGCCAGCTGAGTGGTCAATG
 GCACAACTCGGGCTCACTGCTCTCCGCTCCCGGGTTCAAGTGATTTCTCATGCTCAGCTTCTGAGTAGCTG
 CAGTAGCTGAGATTATAGGCACTACCAACAGCCCGCTCAATTTTGTGTTGTTGTTGAGACAGCTGCTGCTATGTG
 ACCAGGGCTGGAATGATTTCACTGCTCACTGCAACCTTCGCCACCTGGGTTCCAGCAATTCCTGCTCAGCTTCC
 CGAGTAGCTGAGATTATAGGCACTACCAACAGCCCGCTCAATTTTGTGTTGTTGTTGAGACAGCTGCTGCTATGTG
 CATGTTGGCCAGGCTGTGCTGCAACTCCTGACCTTAGGTGATCACTCGCTCTCATCTCCCAAGTGTGGGATTCAGG
 ACAGGCGTGAGCCACCGCTGCTGCTGCCACGCCCACTAATTTTGTATTTTATGAGAGCAGGGTTCCTACCATGT
 TGGCCAGGCTGCTCTTGAACCTCTGACCTCAGGTAACTGACCTGCGCTTCCAAAGTGTGGGATACAGG
 TGTGAGCCACACGCCCGGTACATATTTTTAAATTGAAATCTACTATTTATGTGATCTTTTGGAGCTACAGAG

1015671.22101

[illegible]

FIGURE 180

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492

>subunit 1 of 1, 837 aa, 1 stop

>MW: 90167, pI: 8.39, NX(S/T): 1

MSQTGSHPGRGLAGRWLWGAQPCLLLPIVPLSWLVWLLLLLLASLPSARLASFLPREEEIV
 FPEKLYSGSVLPFGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGGQAPELGGAEF
 GTYLTGTINGDPESVASLHWGGALLGVLYRGAELHLQPLEGGTPNSAGGPGAHILRRKSP
 ASGQGPMCNVKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHAGAGLKRYLLTVMAA
 AAKAFKHPSIRNPVSLVVTRLVILGSGEEGPQVGPSAAQTLRSFCAWQRLNTPEDSGPDHF
 DTAILFTRQDLCGVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHD
 NSKPCISLNGPLSTSRHVMAPVMAHVDPEEPWSPCSARFTDFLDNQYGHCLLDKPEAPLHL
 PVTFFPKDYADARQCQLTFGPDSSRHCPQLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCG
 PAQACMGGRCLEHMDQLQDFNIPQAGGWGPWGPWGDSCRTCGGGVQFSSRDCTRPVPRNGGKY
 CEGRRTRFRSCNTEDECPGTSALTFREEQCAAYNHRDTLFSFPFGPMDWVPRYTGVAPODQCK
 LTCQARALGYVLEPRVVDGTPCSPDSSSVCVQGRCIHAGCDRIIGSKKKFDKCMVCGGDG
 SGCSKQSGSFRKFRYGYNNVVTIPAGATHILVRQGNPQHRSIYLALKLPDGSYALNGEYTL
 MPSPTDVVLPGAVSLRYSGATAAETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPT
 PSTPRPTQDWLHRRQAQILEILRRRPWAGRK

Important features of the protein:

Signal peptide:

amino acids 1-48

N-glycosylation site.

amino acids 68-71

Glycosaminoglycan attachment site

amino acids 188-191, 772-775

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 182-185

Tyrosine kinase phosphorylation site.

amino acids 730-736

N-myristoylation sites.

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-
 172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-
 582, 679-684, 682-687, 763-768

Amidation sites.

amino acids 560-563, 834-837

Leucine zipper pattern.

amino acids 17-38, 24-45

Neutral zinc metalloproteinases, zinc-binding region signature.

amino acids 358-367

FIGURE 181

CAGCAGTGGTCTCTCACTCCTCTCAAAGCAAGGAAAAGAGTACTGTGTGCTGAGAGACC**ATGG**
 CAAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCC
 AAGAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTTTGGTATCCTGGCCCTAACTCT
 AATTGCTCTGTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACA
 TGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGATTACATGGAATTGATCCTGTG
 ACCAGAACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTT
 TAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGA
 TTAAAGTGATTCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACC
 ACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAAGCCTATTGAAAACCGAGA
 TTTTCTTAAAAATCCAAAATTCCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATC
 CCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTT
 CCTGCCAACGAAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGT
 AGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTG
 AAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGT
 CGAGGCAACCGCTATTGCCGCCGCTGTGTGAACCTTTACTAGGCTACTACCCATATCCATA
 CTGTACCAAGGAGGACGAGTCATCTGTCTGTCTCATCATGCCTTGTAAGTGGTGGGTGGCCC
 GCATGCTGGGGAGGGTCT**TAAT**AGGAGGTTTGAGCTCAAATGCTTAAAGTCTGGCAACATAT
 AATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCGGTAGCCAGCT
 CTCAGAAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACC
 AAAAAAAAAAAAAAAAAA

FIGURE 182

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73727
><subunit 1 of 1, 317 aa, 1 stop
><MW: 37130, pI: 5.18, NX(S/T): 3
MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVFGILALTILVLFWGSKHFWEVPPKAY
DMEHTFYSSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKT
QIKVIPEFSEPEEEIDENEEITTTFFEQSVIWWPAEKPIENRDFLKNKILEICDNVTMYWI
NPTLISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDY
TENGIEFDPMLDERGYCCIIYCRGRNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWV
ARMLGRV
```

Important features of the protein:**Signal peptide:**

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 183

GCGGAAC TGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCC
 CGGCTGGCCTAGGCAGGCAGCCGACCC**ATG**GCCAGCACGGCCGTGCAGCTTCTGGGCTTCCT
 GCTCAGCTTCTGGGCATGGTGGGCACGTTGATCACACCATCCTGCCGCACTGGCGGAGGA
 CAGCGCACGTGGGCACCAACATCCTCACGGCCGTGTCTACCTGAAAGGGCTCTGGATGGAG
 TGTGTGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCC
 CCAAGACCTCCAGGCTGCCCCGCGCCCTCATGGTCATCTCTGCTGCTCTCGGGCATAGCCT
 GCGCCTGCGCCGTCTATCGGGATGAAGTGCACGCGCTGCGCCAAGGGCACACCCGCCAAGACC
 ACCTTTGCCATCCTCGGCGGCACCCCTCTTCATCCTGGCCGGCCCTCCTGTGCATGGTGGCCGT
 CTCCTGGACCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCAGCGGCATGA
 AGTTTGAGATTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGT
 GGCACCCTGCTTTGCCTGTCTTCCAGGACGAGGCACCCCTACAGGCCCTACCAGGCCCGGCC
 CAGGGCCACCACGACCACTGCAAAACCCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAG
 ACAATCGGGCCCCCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTG
TGAGTCCCCACAGCCTGCTTCTCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTC
 AATGGAGGCAGGGGTTCCAGCACAAAGTTTACTTCTGGGCAATTTTGTATCCAAGGAAATA
 ATGTGAATGCGAGGAAATGTCTTTAGAGCACAGGGACAGAGGGGGAAATAAGAGGAGGAGAA
 AGCTCTCTATACCAAGACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATAT
 TTATGTGGGTGATTTGATAACAAGTTTAATATAAAGTGACTTGGGAGTTTGGTCAGTGGGGT
 TGGTTTGTGATCCAGGAATAAACCTTGCGGATGTGGCTGTTTATGAAAAAAAAAAAA

FIGURE 184

MASTAVQLLGFLLSFLGMVGTLIITLPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIY
QCQIYRSLLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAGTTPAKTTFAILGGTL
FILAGLLCMVAVSWTTNDVVQNFYNPLLPSGMKFEIGQALYLGFISSSLSLIGGTLLCLSCQ
DEAPYRPYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-103, 115-141, 160-182

FIGURE 185

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCG
 CAGCGCGCAGGCGGGCGGCCAGGATC**ATGT**CCACCACCACATGCCAAGTGGTGGCGTTCCCT
 CCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGATGGACATGTGGAGACCCC
 AGGACCTGTACGACAAACCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGC
 GTGAGGCAGAGTTT**CAGGCTT**CACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGC
 CATGCTGCAGGAGTGCAGGCCCTGATGATCGTAGGCATCGTCTGGTGCCATTGGCCTCC
 TGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCC
 AACATGACACTGACCTCCGGGATCATGTTCTATTGTCTCAGGTCTTTGTGCAATTGTGGAGT
 GTCTGTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCG
 GCATGGGTGGGATGGTGCAGACTGTT**CAGACCAGGTACACATT**TTGGTGCGGCTCTGTTCTGTG
 GGCTGGGTGCGTGGAGGCCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGG
 CCTGGCACCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTG
 CCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACCAAAAAACAAGAAG
 ATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTA
 TGTG**TAA**TGCTCTAAGACCTCTCAGCACGGGCGGAAGAAACTCCCGGAGAGCTCACCCAAAA
 AACAAAGGAGATCCCATCTAGATTTCTTCTTGTCTTTGACTCACAGCTGGAAGTTAGAAAAGC
 CTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCC
 ACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTAT
 TTCTTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCAC
 ATTTTGATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTA
 TTTCCAGCTTATCCCAAGAAACTTTTGAAAGGAAAGAGTAGCCCAAAGATGTTATTTT
 CTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAA
 GAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTG
 TGATCTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTG
 CTGTTGACATCTTCTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAG
 TCCTCTTCTGTCGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAAT
 TTAAGTCCTAAATATAGTTAAATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGA
 AATAGCCTCACCCCTACATGTGGATAGAAGGAAATGAAAAATAATTGCTTTGACATTGTCT
 ATATGGTACTTTGTAAAGTCATGCTTAAAGTACAAATTCATGAAAAGCTCACACCTGTAATC
 CTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTG
 GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGGTG
 GCATACACCTGTAGTCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGG
 AGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCCTCCAGCCAGGTGACATAGCGAGA
 TCCTGTCTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCTTAGGAAGTAGGTTAAA
 ACTAATTCCTTTAA

FIGURE 186

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734
><subunit 1 of 1, 261 aa, 1 stop
><MW: 27856, pI: 8.50, NX(S/T): 1
MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGLRSCVRQSSGFTE
CRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIM
FIVSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTL
IGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTE
DEVQSYPSKHDYV
```

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 187

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCG
 GAGTCCAGCTGGCTAAAACTCATCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTG
 GGCTGTTTCTTGTTGGTGTGGAATGGTGGGCACAGTGGCTGCTACTGTCATGCCTCAGTGG
 AGAGTGTCCGCCITTCATTGAAACAACATCGTGGTTTTTGAAAACTTCTGGGAAGGACTGTG
 GATGAATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGC AAAATCTATGATTTCCCTGCTGG
 CTCTTTCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTC
 TTGGCTTTTCATGATGGCCATCCTTGGCATGAAATGCACCAAGGTGCACGGGGGACAATGAGAA
 GGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGC
 TCATCCCTGTGAGCTGGGTGGCCAAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAAT
 GTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGTCT
 GATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACA
 GATACTCGATACCTTCCCATCGCACACCCAAAAAGTTATCACACCGGAAAGAAGTCACCG
 AGCGTCTACTCCAGAAGTCAGTATGTGTAGTGTGTATGTTTTTTAACTTTACTATAAAGC
 CATGCAAATGACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAAGAACTTTGATTTA
 CTGTTCTTAACTGCCTAATCTTAATTACAGGAACTGTGCATCAGCTATTTATGATTTCTATAA
 GCTATTTTACGAGCAATGAGATATTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAAT
 TTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTTATCATTTTACTTCAAATGACATTGCT
 AAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAG
 TGTAAACATTTATATCTCACATAGAGACATGCTTATATGGTTTTATTTAAAATGAAATGCCAG
 TCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTG
 AAGAAGGTTACTATTAAATTGTTTAAAAACAGCTTAGGGATTAAATGCTCCTCATTATATAATGA
 AGATTA AAATGAAGGCTTTAATCAGCATTGTAAGGAAATGAATGGCTTTCTGATATGCTG
 TTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTCTCCTCCAGAGGCTTTTTTT
 TTCTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTC AAGGGGCTTTGCAATCA
 AACTGCTTTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATG
 GTTTTAGGAAAGTGA AAATATTTTTGTTTTGTATTGTAAGAAGAATGATGCATTTTGACAA
 GAAATCATATATGTATGGATATATTTTAATAAGTATTGAGTACAGACTTTGAGGTTTCATC
 AATATAAATAAAAGAGCAGAAAAATATGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACA
 ACAAAAAAAGTTGTCTTTGAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATTG
 TCATTTTTGTCTGTGAAAAATAAATTTCTTCTTGTACCATTTCTGTTTGTGTTTTACTAAA
 ATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAAACTGACAATCCAATTGGA
 AAGTTTGTGTCGACGCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTTTATACATTTATA
 TTAATAAATTGTACATTTTTCTAATT

FIGURE 188

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73735
><subunit 1 of 1, 225 aa, 1 stop
><MW: 24845, pI: 9.07, NX(S/T): 0
MATHALEIAGLFLGGVGMVGTVAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRM
QCKIYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGI
IFIITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVF
CCNEKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV
```

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 189

TCGCC**ATG**GCCTCTGCCGAATGCAGATCCTGGGAGTCGTCTGACACTGCTGGGCTGGGTG
AATGGCCTGGTCTCCTGTGCCCTGCCCATGTGGAAGGTGACCGCTTTCATCGGCAACAGCAT
CGTGTGGCCAGGTGCTGTGGGAGGGCCTGTGGATGTCTGCGTGGTGCAGAGCACCGGCC
AGATGCAGTGC AAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGGCTGCACGT
GCCCTCTGTGTCA TCGCCCTCCTTGTGGCCCTGTTGGCTTGCTGGTCTACCTTGCTGGGGC
CAAGTGTACCACCTGTGTGGAGGAGAAGGATTCCAAGGCCCGCCTGGTGCTCACCTCTGGGA
TTGTCTTTGTCACTCTAGGGGTCTTGACGCTAATCCCCGTGTGCTGGACGGCGCATGCCATC
ATCCGGGACTTCTATAACCCCTGGTGGCTGAGGCCCAAAGCGGGAGCTGGGGGCCTCCCT
CTACTTGGGCTGGGCGGCCTCAGGCCTTTTGTGTGGGTGGGGGGTTGCTGTGCTGCACTT
GCCCCTCGGGGGGGTCCCAGGGCCCCAGCCATTACATGCCCCGCTACTCAACATCTGCCCT
GCCATCTCTCGGGGGCCCTCTGAGTACCCTACCAAGAATTACGTC**TGA**CGTGGAGGGGAATG
GGGGCTCCGCTGGCGCTAGAGCCATCCAGAAGTGGCAGTGCCCAACAGCTTTGGGATGGGTT
CGTACCTTTTGTCTCTGCCTCCTGCTATTTTCTTTTACTGAGGATATTTAAATTCATTT
GAAAACTGAGCCAAGGTGTTGACTCAGACTCTCACTTAGGCTCTGCTGTTTCTCACCTTGG
ATGATGGAGCCAAAGAGGGGATGCTTTGAGATTCTGGATCTTGACATGCCCATCTTAGAAGC
CAGTCAAGCTATGGAATAATGCGGAGGCTGCTTGCTGTGCTGGCTTTTGCAACAAGACAGAC
TGTCCTCCAAAGAGTTCCTGTGCTGCTGGGGGCTGGGCTTCCCTAGATGTCACTGGACAGCTG
CCCCCATCTACTCAGGTCTCTGGAGCTCCTCTCTTCACCCCGGAAAAACAATCATCTG
TTAACAAAGGACTGCCACCTCCGGAACCTCTGACCTCTGTTTCTCCGTCTGATAAGAGC
TCCACCCCCCAGGGCCAGGTCCAGCTATGTAGACCCCCGCCCCACCTCCAACACTGCACC
CTTCTGCCCTGCCCCCTCGTCTCACCCCTTTTACTCACATTTTTATCAAATAAAGCATG
TTTTGTTAGTGCA

FIGURE 190

```
>>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73736
>>subunit 1 of 1, 220 aa, 1 stop
>>MW: 23292, pI: 8.43, NX(S/T): 0
MASAGMQILGVVLTLLGWVNLVSCALPMWKVTAFIGNSIVVAQVVWEGLWMSCVVQSTGQM
QCKVYDSSLALPQDLQAARALCVIALLVALFQGLLVYLAGAKCTTCVEEKDSKARVLVLTSGIV
FISGVGLTFLIPVCWTAHAIIRDVFYNFLVAEQRELGAASLYLGWAASGLLLGGGLLCTCP
SGGSGQPSHMYRYSTAPATISRGSEYPTKNVL
```

Transmembrane domains:

amino acids 8-30 (type II), 82-102, 121-140, 166-186

FIGURE 191

GCCAAGGAGAACATCATCAAGACTTCTCTAGACTCAAAAGGCTTCCACGTTCTACATCTTG
 AGCATCTTCTACCACTCCGAATTGAACCAGTCTTCAAAGTAAAGGCAATGGCATTTTATCCC
 TTGCAAATTGCTGGGCTGGTTCCTTGGGTTCCCTTGGCATGGTGGGACTCTTGCCACAACCCT
 TCTGCCTCAGTGGTGGAGTATCAGCTTTTGTGGCAGCAACATTATTGTCTTTGAGAGGCTC
 TGGGAAGGGCTCTGGATGAATTGCATCCGACAAGCCAGGGTCCGGTTGCAATGCAAGTTCTA
 TAGCTCCTTGTGGCTCTCCCGCCTGCCCTGGAAACAGCCGGGCCCTCATGTGTGTGGCTG
 TTGCTCTCTCCTTGATCGCCCTGCTTATTGGCATCTGTGGCATGAAGCAGGTCCAGTGCACA
 GGCTCTAACGAGAGGGGCCAAAGCATACCTTCTGGGAACCTCAGGAGTCTCTTCATCCTGAC
 GGGTATCTTCGTTCTGATTCCGGTGAGCTGGACAGCCAATATAATCATCAGAGATTCTACA
 ACCCAGCCATCCACATAGGTCAGAAACGAGAGCTGGGAGCAGCACTTTTCCTTGGCTGGGCA
 AGCGCTGCTGCTCCTTTCATTGGAGGGGGTCTGCTTTGTGGATTTTGTGCTGCAACAGAAA
 GAAGCAAGGGTACAGATATCCAGTGCCTGGCTACCGTGTGCCACACAGATAAGCGAAGAA
 ATACGACAATGCTTAGTAAGACCTCCACCAGTTATGTCTTAATGCCTCCTTTTGGCTCCAAGT
 ATGGACTATGGTCAATGTTTTTTATAAAGTCCTGCTAGAAACTGTAAGTATGTGAGGCAGGA
 GAACCTTGCTTTATGTCTAGATTTACATTGATACGAAAGTTTCAATTTGTTACTGGTGGTAGG
 AATGAAAATGACTTACTTGGACATTCTGACTTCAGGTGTATTAATGCATTGACTATGTGTG
 GACCCAATCGCTGCTCCAATTTTCATATTCTAAATTCAAGTATACCCATAATCATTAGCAAG
 TGTACAATGATGGACTACTTATTACTTTTTGACCATCATGTATTATCTGATAAGAATCTAAA
 GTTGAATTGATATTCTATAACAATAAACATATACCTATTCTA

FIGURE 192

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73737
><subunit 1 of 1, 173 aa, 1 stop
><MW: 18938, pI: 9.99, NX(S/T): 1
MNCIRQARVRLQCKFYSSLLALPPEATRALMCAVALSLIALLLIGICGMKQVQCTGSNER
AKAYLLGTSGVLFILTGIFVLIPVSTANRIIRDIFYNPAIHIGQKRELGAALFLGWASAVAL
FIGGGLLCGFCFCCNRKKQGYRYPVPGYRVPHPTDKRRNTTMSKSTSYV
```

Important features of the protein:

Transmembrane domains:

amino acids 31-51, 71-90, 112-133

N-glycosylation site.

amino acids 161-164

FIGURE 193

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGC**ATGA**AAGATCACTGGGGGT
 CTCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCC
 AAAAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCA
 CATACCTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACCTGTGTACC
 GAGAGCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCT**TAA**ATTCTCCA
 TGGACATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTG
 AGTTTCTTTTCAGTTTCTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTG
 ACTGAATGGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCAT
 TTTTTTTTTAACACGTCAATAAAAAAATAATCTCCAGA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172

FIGURE 194

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73739
```

><subunit 1 of 1, 85 aa, 1 stop

><MW: 9232, pI: 7.94, NX(S/T): 0

MKITGGLLLLCTVVYFCSSSEAAASLPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGN
 ECHLCTESLKSNGRVQFLHDGSC

Signal peptide:

amino acids 1-19

FIGURE 195

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCTCGCCCTCCTGCTCGCGCCC
CGCCGCC**ATGG**GCTGCCTCCCCCGCGCGGCCTGCTGTCTGGCCCTGACCGGGCTGGCGCTGC
TCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAATAAACTCAAGCTGATGCTTCAA
AAACGAGAAGCACCTGTTCCAATAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGA
ATTCTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGACCGGACTCGGCCCGAGGTGC
AGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATC
ACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCA
CTATGATGAAGACTCTGCAATTGGTCCCCGGAGCCCCACGGCTTTAGGCATGGAGCCAGCG
TCAACTACGATGACTACT**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATT
TCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTTCAGCAGAT
CTTTTCTACCTACTTTGTGTGATCAAAAAAGAAGGTTAAACAACACATGTAAATGCCTTT
TGATATTTTCATGGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTAAAAAGA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160

FIGURE 196

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742

><subunit 1 of 1, 148 aa, 1 stop

><MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLLCWGPGGISGNKLLMLQKREAPVPTKTKVAVDENKAKEFL
GSLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRDRNGHEYYGDDYQRHYD
EDSAIGPRSPYGFRHGASVNYDDY

Signal peptide:

amino acids 1-30

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170

FIGURE 197

CGGCTCGAGCCCGCCGGAAGTGCCCGAGGGGCCGCGATGGAGCTGGGGGAGCCGGGCGCTC
 GGTAGCGCGGCGGGCAAGGCAGGCGCC**ATG**ACCCTGATTGAAGGGGTGGGTGATGAGGTGAC
 CGTCTTTTCTCGTGCTTGCTGCCTTCTGGTGCTGGCCCTTGCTGGGTCTCAACGCACA
 CCGCTGAGGGCGGGACCCACTGCCCCAGCCGTAGGGACCCCAACGCCATCCCAGCCCAGC
 GCAGCCATGGCAGCTACCGACAGCATGAGAGGGGAGGCCCCAGGGGCAGAGACCCCCAGCCT
 GAGACACAGAGGTCAAGCTGCACAGCCAGAGCCAGCACGGGGTTCACAGCAACACCGCCAG
 CCCCAGACTCCCCGAGGAGCCCCCTCGTGCTACGGCTGAAATTCTCAATGATTAGAGCAG
 GTGGCCAGGGCCTGGCCCCACGACACCATTGGCTCCTTGAAAAGGACCCAGTTTCCCGGCCG
 GGAACAGCAGGTGCGACTCATCTACCAAGGGCAGCTGCTAGGCGACGACACCCAGACCCCTGG
 GCAGCCTTCACCTCCCTCCCAACTGCGTTCTCCACTGCCACGTGTCCACAGAGTTCGGTCCC
 CCAATCCCCCTGCCCGCGGGGTCCGAGCCCGGCCCTCCGGGCTGGAAATCGGCAGCCT
 GCTGCTGCCCCTGCTGCTCCTGCTGTTGCTGCTGCTCTGGTACTGCCAGATCCAGTACCGGC
 CCTTCTTTCCCTGACCGCCACTCTGGGCCTGGCCGGCTTACCCCTGCTCCTCAGTCTCCTG
 GCCTTTGCCATGTACCGCCCG**TAG**TGCCCTCCGCGGGCGCTTGGCAGCGTCGCCGGCCCCCTCC
 GGACCTTGCTCCCCGCGCCGCGGGGAGCTGCTGCCCTGCCAGGCCCGCTCTCCGGCCTG
 CCTCTTCCCGCTGCCCTGGAGCCCAGCCCTGCCCGCAGAGGACTCCCGGGACTGGCGGAGG
 CCCCGCCCTGCGACCGCGGGGCTCGGGGCCACCTCCCGGGGCTGCTGAACCTCAGCCCGCA
 CTGGGAGTGGGCTCCTCGGGGTGCGGCATCTGCTGTCGCTGCCTCGGCCCGGGCAGAGCCG
 GGCGCCCCGGGGGCCCGTCTTAGTGTTCTGCCGGAGGACCCAGCCGCTCCAATCCCTGAC
 AGCTCCTTGGGCTGAGTTGGGGACGCCAGGTGGTGGGAGGCTGGTGAAGGGGAGCGGGGAG
 GGGCAGAGGAGTTCCCCGGAACCCGTGCAGATTAAAGTAACTGTGAAGTTTTAAAAAAAAA
 AAAAAAA

FIGURE 198

MTLIEGVGDEVTVLFSVLACLLVLALAWVSTHTAEGGDPLFPQPSGTPTPSQPSAAMAATDSM
RGEAPGAETPSLRHRGQAAQPEPSTGFTATPPAPDSPQEPLVLRKFLNDSEQVARAWPHDT
IGSLKRTQFPGREQQVRLIYQGQLLGDDTQTLGSLHLPPNCVVLHCHVSTRVGPPNPFCPPGS
EPGPSGLEIGSLLLPLLLLLLLLLLWYCQIQYRPFPLTATLGLAGFTLLLSLLAFAMYRP

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 195-217

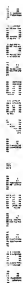


FIGURE 199

GAGATTGGAAACAGCCAGGTTGGAGCAGTGAGTGAGTAAGGAAACCTGGCTGCCCTCTCCAG
 ATTCCCCAGGCTCTCAGAGAAGATCAGCAGAAAGTCTGCAAGACCCTAAGAACCATCAGCCC
 TCAGCTGCACCTCCTCCCCCTCAAGG**ATG**ACAAAGGCGCTACTCATCTATTTGGTCAGCAGC
 TTTCTTGCCCTAAATCAGGCCAGCCTCATCAGTCGCTGTGACTTGCCCCAGGTGCTGCAGCT
 GGAGGACTTGGATGGGTTTGAGGGTTACTCCCTGAGTGACTGGCTGTGCCTGGCTTTTGTGG
 AAAGCAAGTTCAACATATCAAAGATAAATGAAAATGCGGATGGAAGCTTTGACTATGECCTC
 TTCCAGATCAACAGCCACTACTGGTGCAACGATTATAAGAGTTACTCGGAAAACCTTTGCCA
 CGTAGACTGTCAAGATCTGCTGAATCCCAACCTTCTTGCAAGGCATCCACTGCGCAAAAAGGA
 TTGTGTCCGGAGCACGGGGGATGAACAACCTGGGTAGAATGGAGGTTGCACTGTTTCAGGCCGG
 CCACTCTCCTACTGGCTGACAGGATGCCGCCTGAGAT**TGA**AACAGGGTGCGGGTGCAACCGTGG
 AGTCATTCCAAGACTCCTGTCTCACTCAGGGATTCTTCATTTCTTCTTCTACTGCCTCCA
 CTTTCATGTTATTTTCTTCCCTTCCCATTTACAACATAAACTGACCAGAGCCCCAGGAATAAA
 TGGTTTTCTTGGCTTCTCCTTACTCCCATCTGGACCCAGTCCCCCTGGTTCTGTCTGTTAT
 TTGTAAACTGAGGACCACAATAAGAAATCTTTATATTTATCG

FIGURE 200

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746

><subunit 1 of 1, 148 aa, 1 stop

><MW: 16896, pI: 6.05, NX(S/T): 1

MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEYGYSLSDWLCLAFVESKFNISKI
NENADGSFDYGLFQINSHYWCNDYKSYSENLC HVDCQDLLNP NLLAGIHC AKRIVSGARGMN
NWVEWRLHCSGRPLSYWLTGCRLR

Signal peptide:

amino acids 1-18

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

FIGURE 201

TCTGACCTGACTGGAAGCGTCCAAAGAGGGACGGCTGTCAGCCCTGCTTGACTGAGAACCCA
 CCAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGGAAAGAACCTGAG
 CAGAATGGAATCATATATTTTTTCCCAAGGAGAAAACCGGGTAAAGGGAGGGAAGCAATT
 AATTTGAAGTCCCTGTGTAATGGGCTTTCAGAAGGCAATTAAGAATCCACTCAGAGAGGAC
 TTGGGGTGAAACTTTGGGTCTGTGGTTTTCTGATTGTAAGTGGAAGCAGGTCTTGACACAGC
 TGTGTGGCAAATGTCAGGACCAGGTTAAGTGACTGGCAGAAAAAATCCAGGTGGAACAAGCA
 ACCCATGTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGGAGAAAGCTAACTTGAACATGAC
 CTGTTGCATTTGGCAAGTTCTAGCAACATGCTCCTAAGGAAGCGATACAGGCACAGACCATG
 CAGACTCCAGTTCTCCTGCTGCTCCTGATGCTGGGATGCGTCCTGATGATGGTGGCGATGT
 TGCACCTCCCCACCACACCTGCAACAGACTGTCACAGCCCAAGCCAGCAAGCACAGCCCT
 GAAGCCAGGTACCGCTGGACTTTGGGGAATCCAGGATTGGGTACTGGAAGCTGAGGATGA
 GGGTGAAGACTACAGCCCTCTGGAGGGCTGCCACCCTTTATCTACTGCGGGAGGATCAGC
 TGCTGGTGGCCGTGGCCTTACCCCAGGCCAGAAGGAACAGAGCCAGGGCAGGAGAGGTGGG
 AGCTACCCGCTCATCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGACTGGGG
 GGCTGATGAGGACGGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTAGCCCTGGACCCAC
 GTGGCCTCCAGGAGGCCTCAGTGCCCGCATCCCCCTCCAGAGGGCTCTGCCCGAGGTGCGG
 CACCCACTGTGTCTGCAGCAGCACCTCAGGACAGCCTGCCACAGCCACGCCTCATCTCTGT
 TTTCCATGATGAGGCTGGTCCACTCTCCTGCGGACTGTACACAGCATCTCGACACAGTGCC
 CCAGGGCCTTCTGAAAGGAGATCATCTCTGCGGACGACCTCAGCCAGCAAGGCAACTCAAG
 TCTGCTCTCAGCGAATATGTGGCCAGGCTGGAGGGGTGAAGTTACTCAGGAGCAACAAGAG
 GCTGGGTGCCATCAGGGCCCGGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCT
 TCATGGATGCCCACTCCGAGTGCCACCCAGGCTGGCTGGAGCCCTCCTCAGCAGAATAGCT
 GGTGACAGGAGCCGAGTGGTATCTCCGGTGATAGATGTGATTGACTGGAAGACTTCCAGTA
 TTACCCCTCAAAGCACTGCAGCGTGGGTGTTGGACTGGAAGCTGGATTCCACTGGGAAC
 CTTTGGCAGAGCATGTGAGGAAGGCCCTCCAGTCCCCCATAAGCCCCCTCAGGAGCCCTGTG
 GTGCCCGGAGAGGTGGTGGCCATGGACAGACATTACTTCCAAAACACTGGAGCGTATGACTC
 TCTTATGTGCTGCGAGGTGGTGAACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTG
 GCTCTGTTGAAATCCTTCCCTGCTCTCGGGTAGGACACATCTACCAAAATCAGGATTCCTAT
 TCCCCCTCGACCAGGAGGCCACCTTGAGGAACAGGGTTCCGATTGCTGAGACCTGGCTGGG
 GTCATTCAAAGAAACCTTACAAGCATAGCCAGAGGCCCTTCTCCTTGAGCAAGGCTGAGA
 AGCCAGACTGCATGGAACGCTTGCAGCTGCAAAGGAGACTGGGTTGTGCGACATTCACACTGG
 TTTCTGGCTAATGTCTACCCTGAGCTGTACCCATCTGAACCCAGGCCAGTTTCTCTGGAAA
 GCTCCACAACACTGGACTTTGGGCTCTGTGCAGACTGCCAGGCAGAAAGGGACATCTGGGCT
 GTCCTATGTTGTTGGCTCCTTGCAGTGACAGCCGCGCAGCAACAGTACCTGTCAGCACACCAG
 AGGAAGGAGATTCACTTTGGCAGGCCACAGCACTGTGCTTTGCTGTGAGGCAGGAGCAGCT
 GATTCTTCAGAACTGCACGGAGGAAGGCCCTGGCCATCCACAGCAGCACTGGGACTTCCAGG
 AGAATGGGATGATTGTCCACATCTTTCTGGGAATGCATGGAAGCTGTGGTGAAGAAAAC
 AATAAGATTGTACCTGCGTCCGTGTGATGGAAGAACCCGCGCAGGTCAGGATTTGACCA
 GATAAATGCTGTGGATGAACGATGAATGTCAATGTGAGAAGGAAAAGAAATTTGGCCATC
 AAAATCAGCTCCAAGTGAACGTAAAGAGCTTATATTTCTCATGAAGCTGATCCTTTGTGT
 GTGTGCTCCTGTGTGTAGGAGAGAAAAAGCTCTATGAAAGATATAGGAAGTTTCTCCTTT
 TCACACCTTATTTCAATGTACTGCTGGCTGCTTA

10015671-721101

FIGURE 202

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760

><subunit 1 of 1, 639 aa, 1 stop

><MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFG
ESQDWVLEAEDEGEEYSPLEGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPR
RQDKEAPKRDWGADEDEGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQQHP
QDSLPTASVILCFHDEAWSTLLRTVHSILDTVPRAFLKEIILVDDLSQQGQLKSALSEYVAR
LEGVKLLRSNKRIGAIRARMLGATRATGDVLVFMDAHCECHPGWLEPLLSRIAGDRSRVVSF
VIDVIDWKTQYYPSKDLQRGVLDWKLDHFHWEPLPEHVRKALQSPISPIRSFVVPGEVVAMD
RHYFQNTGAYDSLMSLRGGENLELSFKAWLCGGGSVEILPCSRVGHYQNDSSHSPLDQEATL
RNRVRIAETWLGSEFKETFYKHSPEAFSLSKAEKPDCMERLQLQRRLGCRTFHWFLANVYPEL
YPSEPRPSFSGKLHNTGLGLCADCAEGDILGCFMVLAPCSDSRQQQYLQHTSRKEIHFGSP
QHLCEFAVRQEQVILQNCTEEGLAIHQHWFQENGMIVHILSGKCMNAVQENNKDLYLRPC
DGKARQQWRFDQINAVDER

Signal peptide:

amino acids 1-28

FIGURE 203

CGCCAAGCATGCAGTAAAGGCTGAAAACTGGGTACACAGCTGAGGAAGACCTCAGACATGGA
 GTCCAGGATGTGGCCTGCGCTGCTGCTGCCACCTCCTCCCTCTCTGGCCACTGCTGTTTGC
 TGCCCCCTCCCACCGCCTGCTCAGGGCTCTTCATCCTCCCCTCGAACCCCAACGACCCAGCC
 CGCCCCCGTGTGCCAGGGGAGGCCCTCGGCCACAGTCATGTGTGCGTGTGGGAGCGAGC
 ACCTCCACCAAGCCGATCTCCTCGGGTCCCAAGATCACGTGCGGCAAGTCTCTGCTGGCACTG
 CACCCCAAGCCACCCATCAGGCTTTGAGGAGGGGCCGCCCTCATCCCAATACCCCTGGGCT
 ATCGTGTGGGGTCCCACCGTGTCTCGAGAGGATGGAAGGGGACCCCAACTCTGCCAATCCCGG
 ATTTCTGGACTATGGTTTTGCGACCCCTCATGGGCTCGCAACCCCAACCCCAACTCAGACT
 CCATGCGAGGTGATGGAGATGGGCTTATCCTTGAGAGGGCACCTGCCACCCCTGCGGCCATTC
 CTGTTGCGGGGGCGTGGGGAAGGTGTGGACCCCAAGCTCTATGTACAATTACCATCTCCAT
 CATCATTTGTTCTCGTGGCCACTGGCATCATCTTCAAGTTCTGCTGGGACCGCAGCCAGAAGC
 GACGCAGACCCCTCAGGGCAGCAAGGTGCCCTGAGGCAGGAGGAGAGCCAGCAGCCACTGACA
 GACCTGTCCCCGGCTGGAGTCACGTGTGCTGGGGGCTTCGGGGGACTCACCTACCCCCACCCC
 TGACCATGAGGAGCCCCGAGGGGACCCCGGCTGGGATGCCCCACCCCAAGGGGGCTCCAG
 CCTTCCAGTTGAACCGGTGAGGGCAGGGGCAATGGGATGGGAGGGCAAGAGGGAAGGCCAAC
 TTAGGTCTTCAGAGCTGGGGTGGGGGTGCCCTCTGGATGGGTAGTGAGGAGGCAGGCGTGGC
 CTCACACAGCCCTTGCCCTCCCAAGGGGCTGGACACAGCTCCTCTCTGGGAGGCACCCCTTC
 CTTCTCCAGTCTCTCAGGATCTGTGTCTTATCTCTGCTGCCATAACTCCAACCTCTGCCCT
 TCTTTGGTTTTTTCTCATGCCACCTTGCTCTAAGACAACCTTGCCCTCTTAACCTTGATTCCC
 CCTCTTTGCTTTGAACCTTCCCCTTCTATTCTGGCCTACCCCTTGGTTCTTGACTGTGCCCTT
 TCCCCTTCTCTCAGGATTTCCCCTTGGTGAATCTGTGATGCCCCCAATGTTGGGGTGCAGCC
 AAGCAGGAGGCCAAGGGGCGGACACAGCCCCCATCCCACTGAGGGTGGGGCAGCTGTGGGA
 GCTGGGGCCACAGGGGCTCTGCGCTCTGCCCTTGACACACCCGGAACACTCCCAGCC
 CCACGGGCAATCCTATCTGCTCGCCCTCCTGCAGGTGGGGGCTCACATATCTGTGACTTCG
 GGTCCCTGTCCCCACCCCTTGTCACCTCACATGAAAGCCTTGACACACTCACCTCCACCTTAC
 AGGCCATTGTGCACACGCTCTGACCCCTCTCCCCGTCCATACCGCTCCGCTCAGCTGACTCT
 CATGTCTCTCGTCTCACATTTGCACTCTCTCCTTCCACATTTCTGTGCTCAGCTCACTCAG
 TGGTCAGCGTTTCTGCACACTTTACCTCTCATGTGCGTTTTCCGGCCTGATGTTGTGGTGG
 TGTGCGGGGTGCTCACTCTCTCCCTCATGAACACCCACCCACCTCGTTTTCCGGAGCCCCCTGC
 GTGCTGCTCCAGAGGTGGGTGGGAGGTGAGCTGGGGGCTCCTTGGGCCCTCATCGGTGATGG
 TCTCGTCCCATTCCACACCATTTGTTTCTCTGTCTCCCCATCCCTACTCCAAGGATGCCGGCA
 TCACCTGAGGGCTCCCCCTTGGGAATGGGGTAGTGAGGCCCCAGACTTCACCCCAAGCCCA
 CTGCTAAAAATCTGTTTTCTGACAGATGGGTTTTGGGGAGTCGCTGCTGCACCTACATGAGAA
 AGGGACTCCCAATTTGCCCTTCCCCTTCTCCTACAGTCCCCTTTTGTCTGTCTGTCTCGGCTG
 TCTGTGTGTGCGCATTTCTTGACTTCAGAGCCCCCTGAGCCAGTCCCTCCCTTCCCTAGCCCT
 CCCCTTGGGCTCCCTAACTCCACCTAGGCTGCCAGGGACCGAGTCAGCTGTTTCAAGGCC
 ATCGGGAGCTCTGCCTCAAGTCTACCCCTTCCCTTCCCGGACTCCCTCCTGTCCCCCTCCTTT
 CCTCCCTCCTTCCCTTCCACTCTCCTTCTCTTTTGGCTTCCCTGCCCCCTTCCCCCTCCTCAGGTT
 CTTCCCTCCTTCTCATGGTTTTTCCACCTTCTCCTTCCCTTCTTCCCTGGCTCCTAGGCT
 GTGATATATATTTTGTATTATCTCTTTCTCTTCTTGGTGTATCATCTTGAATTACTGTG
 GGATGTAAGTTTCAAAATTTTCAATAAAGCCTTTTGAAGATAA

FIGURE 204

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPGAV
PGRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 205

GTTAACCAGCGCAGTCCTCCGTGCGTCCCGCCCGCCGCTGCCCTCACTCCCGGCCAGG**ATGG**
 CATCCTGTCTGGCCCTGCGCATGGCGCTGCTGCTGGTCTCCGGGGTTCTGGCCCTGCGGTG
 CTCACAGACGATGTTCCACAGGAGCCCGTGCCACGCTGTGGAACGAGCCGGCCGAGCTGCC
 GTCGGGAGAAGGCCCCGTGGAGAGCACACGCCCCGGCCGGGAGCCCGTGGACACCGGTCCCC
 CAGCCCCACCGTCGCGCCAGGACCCGAGGACAGCACCGCGCAGGAGCGGCTGGACCAGGGC
 GCGGGTCTGCTGGGGCCCGCGCTATCGCGGCCATCGTGATCGCCGCCCTGCTGGCCACCTG
 CGTGGTGTGCGCTCGTGGTCTGTCGCTGAGAAAGTTTTCTGCCTCCT**TGA**AGCGAATAAA
 GGGGCCGCGCCCGCCGCGCGCGACTCGGCAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177

FIGURE 206

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398
><subunit 1 of 1, 121 aa, 1 stop
><MW: 12073, pI: 4.11, NX(S/T): 0
MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVESTSPGREPVDTG
PPAPTIVAPGPEDESTAQERLDQGGGSLGPGATAAIVIAALLATCVVLALVVVALRKFSAS
```

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 91-110

Glycosaminoglycan attachment site.

amino acids 44-47

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 116-119

N-myristoylation site.

amino acids 91-96

FIGURE 207

GGCCGTTGGTTGGTGCGCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGG
 CGGGCCGGGACGGGC**ATG**GCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGCGCTGGCCCA
 CGGCTGTCTGCACTGCCACAGCACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGA
 ACTTCAAGTCTCTGGTGGGTGGCGACATCCCCGTGTAGGGGCGCTGCTCACCGACTGGAGC
 GACGACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGAGAAGCTGGA
 CCAAGTGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCC
 CCGGTATTTCCCCAACGAGTGCGAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAAC
 GCCATCATCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGG
 ACCCAGCCTAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCC**TA**AGTAGCCC
 CCAGAGGCGCTGGGAGTGTTGCCACCGCCCTCCCTGAAGTTTGCTCCATCTCACGTGGGG
 GTCAACCTGGGGACCCCTTCCCTCCGGGCCATGGACACACATACATGAAAACAGGCCGCAT
 CGACTGTGACACCGCTGTGGCATCTTCCAGTACGAGACCATCTCCTGCAACAACCTGCACAG
 ACTCGCAGTGCCTGCTTTGGCTATACTGCGAGTAGGGCTCAGGCATCACACCCACCCGT
 GCCAGGGCCCTACTGTCCCTGGGGTCCCAGGCTCTCCTTGGAGGGGGCTCCCCGCCTTCCAC
 CTGGCTGTATCGGGTAGGGCGGGCCGTGGGTTCAAGGGCGCACCACTTCCAAGCCTGTGT
 CCCACAGGTCCTCGGCGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACACATAAATAAC
 TGGCACAAGTAAGTCCCTCCTCAAACCAACACAGGCAGTGTGTATGTGAGCACCTCGTG
 GGTGAGTATGTGTGGGCGACAGGCTGGCTCCCTCAGCTCCACGTCCTAGAGGGGCTCCCGA
 GGAGGTGGAACCTCAACCCAGCTCTGCGCAGGAGGCGGCTGCAGTCTTTTCTCCCTCAAAG
 GTCTCCGACCTCAGCTGGAGGCGGGCATCTTTCCTAAAGGGTCCCATAGGGTCTGGTTCC
 ACCCATCCAGGTCTGTGGTCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCC
 ATGGAGGGGCTGACTGCCCCACATTGCCTTTCAGACAGGACACGAGCATGAGGTAAGGCCGC
 CCTGACCTGGACTTCAGGGGGAGGGGGTAAAGGGAGAGAGAGGGGGGCTAGGGGGTCTCT
 AGATCAGTGGGGGCACTGCAGGTGGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTAC
 CTCTGCAACCACACCCATGTGGTGGTTTATGAACAGACCACGCTCCTCTGCCTTCTCCTGG
 CCTGGGACACACAGAGCCACCCGGCCTTGTGAGTGACCCAGAGAAGGGAGGCGCTCGGGAGA
 AGGGGTGCTCGTAAGCCAACACCAGCGTGCCGCGGCCTGCACACCTTCGGACATCCCAGGC
 ACGAGGGTGTGCTGGATGTGGCCACACATAGGACCACAGTCCCAGCTGGGAGGAGAGGCCT
 GGGGCCCCAGGGAGGGAGGCAGGGGGTGGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCC
 CCGCAGCCTGGTATCGCCAGCCTTAAGGTGTCTGGAGCCCCACACTTGGCCAACTGACCT
 TGGAAGATGCTGCTGAGTGTCTCAAGCAGCACTGACAGCAGCTGGGCCTGCCCCAGGGCAAC
 GTGGGGGCGGAGACTCAGCTGGACAGCCCTGCCTGTCACTCTGGAGCTGGGCTGCTGCTGC
 CTCAGGACCCCTCTCCGACCCGGACAGAGCTGAGCTGGCCAGGGCCAGGAGGGCGGGAGG
 GAGGAATGGGGTGGGCTGTGCGCAGCATCAGCGCCTGGGCAGGTCCGAGAGCTGCGGGA
 TGTGATTAAAGTCCCTGATGTTTCTC

FIGURE 208

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399
><subunit 1 of 1, 157 aa, 1 stop
><MW: 17681, pI: 7.65, NX(S/T): 1
MALLLCLVCLTAALAHGCLHCHSNFSKKFSFYRHHVNFKSWWVGDI PVSGALLTDWSDDTMK
ELHLAIPAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAIER
HLAPGSWGGQLSREGPSLAPEGSMPSPRGDL P
```

Signal peptide:

amino acids 1-15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 209

AGCAGGAGCAGGAGAGGGACAATGGAAGCTGCCCCGTCCAGGTTTCATGTTCTCTTATTTCT
 CCTCACGTGTGAGCTGGCTGCAGAAAGTTGCTGCAGAAAGTTGAGAAATCCTCAGATGGTCCTG
 GTGTGCCCAGGAACCCACGTGGCTCACAGATGTCCCAGCTGCCATGGAATTCATTGCTGCC
 ACTGAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCA
 TAGCATGTTGCAAAAATCCCAGGCGTGTCAATTGGGATCAGCACTGATTCTGAGGTTCTGA
 CACACTACAACATCACTGGGAACACCATCTGCCTCTTTGCGCTGGTAGACAATGAACAATG
 AATTTAGAGGACGAAGACATTGAAAGCATTGATGCCACCAAATGAGCCGTTTCATTGAGAT
 CAACAGCCTCCACATGGTGACAGAGTACAACCCTGTGACTGTGATTGGGTTATTCAACAGCG
 TAATTCAGATTTCATCTCCTCCTGATAATGAACAAGGCCCTCCCCAGAGTATGAAGAGAACATG
 CACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAG
 TGGTATGAAAGAAAATGGGAAGGTGATATCATTTTTCAAATAAGGAGTCTCAACTGCCAG
 CTTTGCCAATTTACCAGACTCTAGATGACGAGTGGGATACACTGCCACAGCAGAAGTTTCC
 GTAGAGCATGTGCAAACTTTTGTGATGGATTCTTAAGTGAAAAATTGTTGAAAGAAAATCG
 TGAATCAGAAGGAAAGACTCCAAAGGTGGAACCTCTGACTTCTCCTTGGAACTACATATGGCC
 AAGTATCTACTTTATGCAAAGTAAAAAGGCACAACCTCAAATCTCAGAGACACTAACAACAG
 GATCACTAGGCCTGCCAACCACACACACACGACGTGCACACACGCACGCACGCGTGCACAC
 ACACACGCGCACACACACACACACAGAGCTTCATTTCTGTCTTAAATCTCGTTTTCTC
 TTCTTCCTTCTTTTAAATTTTCATATCCTCACTCCCTATCCAATTTCTTCTTATCGTGCATT
 CATACTCTGTAAGCCATCTGTAAACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATG
 CCTCTATGAAAGAGAGGCATTCTTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGT
 TTGTATACTGCACATGACTTACACACAACATAGTTCCTGCTCTTTTAAGGTTACCTAAGGGT
 TGAACTCTACCTTCTTTATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAGG
 ATGGTTTTAAACACCTTTGTGAAATTGTCTTTTGGCAGAAGTTAAAGGCTGTCTCCAAGTC
 CCTGAACCTCAGCAGAAATAGACCATGTGAAAACCTCATGCTTGGTTAGCATCTCCAACCTCC
 TATGTAAATCAACAACCTGCATAATAATAAAAGGCAATCATGTTATA

FIGURE 210

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30480, pI: 4.60, NX(S/T): 1

MEAPSRFMFLFLLTCELAEEVAAEVEKSSDGPAAQEPTWLTDPAAEFIAATEVAVIG
FFQDLEIPAVPILHSMVQKFPGVSFGISDSEVLTHYNITGNTICLFRDVDNEQLNLEDEDI
ESIDATKLSRFIEINSLHMTVEYNPVTVIGLFNSVIQIHLLLIMNKASPEYEENMHRYQKAA
KLFQGGKILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDDEWDTLPATAEVSVHEVQNF
CDGFLSGKLLKENRESEGKTPKVEL

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 143-162

FIGURE 211

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGGTG
 GCGGAGAGATCAGAAGCCTCTTCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCTCAGG
 GACGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACCATGGCAGCGTCCGCCGGAGCCGGG
 GCGGTGATTGCAGCCCCAGACAGCCGCGCTGGCTGTGGTGGTGCTGGCGGCGGCGCTTGG
 GCTCTTGACAGCTGGAGTATCAGCCTTGAAGTATATACGCCAAAAGAAATCTTCGTGGCAA
 ATGGTACACAAGGGAAGCTGACCTGCAAGTTCAGTCTACTAGTACGACTGGCGGGTTGACC
 TCAGTCTCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTCGTTTTTCCACTACTC
 CCAAGGGCAAGTGTAACCTTGGGAATTATCCACCATTAAAGACAGAATCAGCTGGGCTGGAG
 ACCTTGACAAGAAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACACAATGGCACC
 TATATCTGTGATGTCAAAAACCCCTCCTGACATCGTTGTCCAGCCTGGACACATTAGGCCTCTA
 TGTCGTAGAAAAAGAGAATTGCTGTGTTTCCAGTTTGGGTAGTGGTGGGCATAGTTACTG
 CTGTGGTCTTAGGTCTACTCTGCTCATCAGCATGATTCGGCTGTCCTCTATAGAAGGAAA
 AACTCTAAACGGGATTACACTGGCTGCAGTACATCAGAGAGTTGTCAACAGTTAAGCAGGC
 TCCTCGGAAGTCCCCCTCCGACACTGAGGGTCTTGTAAGAGTCTGCCTTCTGGATCTCACC
 AGGGCCCACTCATATATGCACAGTTAGACCACTCCGGCGGACATCACAGTGACAGAGATTAAC
 AAGTCAGAGTCTGTGGTGTATGCGGATATCCGAAAGAATTAAAGAGAATACCTAGAACATATC
 CTCAGCAAGAAACAAAACCAAACCTGGACTCTCGTGCAGAAAATGTAGCCCATTACCACATGT
 AGCCTTGGAGACCCAGGCAAGGACAAGTACACGTGTACTCACAGAGGGAGAGAAAGATGTGT
 ACAAGGATATGTATAAATATCTATTTAGTCATCCTGATATGAGGAGCCAGTGTGTCATGA
 TGAAAAGATGGTATGATTCTACATATGTACCCATTGCTCTGCTGTTTTTGTACTTCTTTTC
 AGGTCATTTACAATTGGGAGATTTCAGAAACATTCCCTTTCACCATCATTTAGAAATGGTTTG
 CCTTAATGGAGACAATAGCAGATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAATCTAAG
 GGCTTAAGACTGATTAGTCTTAGCATTTACTGTAGTTGGAGGATGGAGATGCTATGATGGAA
 GCATACCCAGGGTGGCCTTTAGCACAGTATCAGTACCATTTATTTGTCTGCCGCTTTTAAAA
 AATACCCATTGGCTATGCCACTTGAAACAAATTTGAGAAGTTTTTTTGAAGTTTTTCTCACT
 AAAATATGGGGCAATTGTTAGCCTTACATGTTGTTGAGACTTACTTTAAGTTTGCACCCCTTG
 AATGTGTGCATATCAATTTCTGGATTCATAATAGCAAGATTAGCAAAGGATAAATGCCGAAG
 GTCACCTTCATCTCGGACACAGTTGGATCAATACTGATTAAGTAGAAAAATCCAAGCTTTGCTT
 GAGAACTTTTGTAACGTGGAGAGTAAAAAGTATCGGTTTTTA

FIGURE 212

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76510
><subunit 1 of 1, 269 aa, 1 stop
><MW: 29082, pI: 9.02, NX(S/T): 3
MAASAGAGAVIAAPDSRRWLWSVLAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKS
TSTTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKSDASINIE
MQFIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLTLLISMI
LAVLYRRKNSKRDTGCTSESLSVPKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSG
GHHSDKINKSESVVYADIRKN
```

Signal peptide:

amino acids 1-37

Transmembrane domain:

amino acids 161-183

FIGURE 213

GCCGGCTGTGCAGAGACGCC**ATG**TACCGGCTCCTGTGTCAGCAGTGA**CT**GCCCGGGCTGCCGCC
 CCCGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGCTGCCGCC
 TCTCGGCCACGGCTGGGTGCGGGGCTCGGGCTGGGGCTGGGGCTCGGGGTGAAGC
 TGGCAGGTGGGCTGAGGGGCGCGGCCCGGCGCAGTCCCCCGGGCCCCGACCTGAGGCG
 TCGCCTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGC
 GCCGCCCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGA
 TCAAGGATGAGGTGGGCGCACCGGGCATAGTGGTTGGAGTTTCTGTAGATGGAAGAAGTC
 TGGTCAGAAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGT
 TATGCGAATTGCTAGCATCAGCAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAG
 CAGGGAACTGGATCTTGATATTCAGTACAACATTATGTTCCCGAATTCCCGAAAAAGAA
 TATGAAGGTGAAAAGGTTTCTGTGACACAAGATTACTGATTTCCCATTTAAGTGGAATTCG
 TCATTATGAAAAGGACATAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGA
 TGAAGAGAAATGTTGCATTTGAGCAAGAAAAAGAGGCAAAAGTAATGAAAAGAATGATTTT
 ACTAAATTTAAACAGAGCAGGAGAATGAAGCCAAATGCCGGAATCAAACCTGGCAAGAA
 AAAGAATGATTTTGAACAAGGCGAATTATATTTGAGAGAAAAGTTTGAAAATTC AATTGAAT
 CCCTAAGATTATTTAAAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTGTGATTCA
 ACTTTTGGCTATACCTACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGA
 CTATATGCAGAAAATATTCATGACTTGGATATGCTGACGACTGTGCAGGAAGAAAACGAGC
 CAGTGATTTACAATAGAGCAAGG**TAA**ATGAATACCTTCTGCTGTGCTAGCTATATCGCATC
 TTAACACTATTTTATTAATTTAAAGTCAAATTTCTTTGTTTCCATTCCAAAATCAACCTGC
 CACATTTTGGGAGCTTTTCTACATGTCTGTTTCTCATCTGTAAAGTGAAGGAAGTAAAACA
 TGTTTATAAAGTAAAAAAA


```
></usr/seqdb2/ssc/DNA/Dnaseqs.min/ss.DNA76522
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41221, pI: 8.54, NX(S/T): 0
MYRLLSAVTARAAAPGGLASSCGRRGVHQRAGLPPLGHGWVGGGLGLGLALGVKLAGGLRG
AAPAQSPAAPDPEASPLAEPPEQESLAPWSPQTAPPCCSRCFARAIESSRDLLHRIKDEVGA
PGIVVGVSVDGKEVWSEGLGYADVENRVCKPETVMRIASISKSLTMVALAKLWEAGKLDLD
IPVQHYVPEFPEKEYEYEGKVSVTTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAF
EQEKEGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKN
DPLFFKPGSQFLYSTFGYTTLLAAIVERASGCKYLDYMQKIFHDLMDLTTVOEENEPVIYNRA
```

amino acids 1-19

amino acids 39-60

STGACACTATAGAAGAGCTATGACGTGCGATGCACGCGTAGCTAAGCTCGGAATTCGGCTCG
AGGCTGGTGGGAAGAAGCCGACATGCGCGCAGCCAGCGCTGGGGCAACCCGGCTGCTCTGCG
TCTTGCTGATGGCGGTAGCAGCGCCAGTCGAGCCCGGGCAGCGGCTGCCGGGCCGGGACT
GGTGCGCAGAGGGCTGGGGCGGAAGGTCGAGAGGGCAGGCCCTGTGGCACGCTGGGGCTGCT
GCTGGAGCACTCATTGAGATCGATGACAGTGCCAACTTCCGAAGCGGGGCTCACTGCTCT
GGAACCAAGCAGGATGGTACCTTGTCCCTGTACAGCGGCAGCTCAGCGAGGAGGAGCGGGGC
CGACTCCGGGATGTGGCAGCCCTGAATGGCTGTACCGGGTCCGGATCCCAGGCACCCGG
GGCCTGGATGGCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGG
TGGAGTCGCACCTGTCCGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGC
GTGTCCGGTGGTGACGCACCCCGGGGGCTGCCGGGGCCATGAGGTGGAGGACGTGGACCTGGA
GCTGTTCAACACCTCGGTGCAGCTGCAGCCGCCACACAGCCCCAGGCCCTGAGACGGCGG
CCTTCATTGAGCGCCTGGAGATGGAACAGGCCAGAGAAGGCCAAGAACCCCCAGGAGCAGAAG
TCCTTCTTCGCCAAATACTGGATGTACATCATTCCTCGTCTGTTCTCTCATGATGTGAG
AGCGCCAGACACCGGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCC
TTTGCTGTGTGCCACCTCCTGTAAAGTCTATTTAAAAACATCGACGATACATTGAAATGTG
TGAACGTTTTGAAAGGCTACAGCTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGG
TCCGTGATGTACAAAGTGTATTGAAATCTAGCTCACTTGATACGTTATTGAAACCCAAAG
GAATGGCTGTCCCATCCTCATGTGGCTGTGTGAGCTCAGCTGTGTTGTGTGGCAGTTTAT
TAAACTGTCCCCAGATCGACACGCAAAAAAAAAA

FIGURE 216

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76529
><subunit 1 of 1, 269 aa, 1 stop
><MW: 28004, pI: 5.80, NX(S/T): 1
MAAASAGATRLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEI
DDSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVRIPRRPGALDGLEA
GGYVSSFVPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELFNTSVQ
LQPPTTAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYWMYIIPVVLFLMMSGAPDTGGQ
GGGGGGGGGGGSGLCVPPSL
```

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 226-243

FIGURE 217

GGAGCGCTGCTGGAACCCGAGCCGGAGCCGAGCCACAGCGGGGAGGGTGGCCTGGCGGCCT
 GGAGCCGGACGTGTCCGGGGCGTCCCGCAGACCGGGGCAGCAGGTCGTCCGGGGGCCACC
ATGCTGGTGACTGCCTACCTTGCTTTGTAGGCCTCCTGGCCTCCTGCCTGGGGCTGGAAC
 GTCAAGATGCCGGGCTAAACCCCTGGAAGGGCCTGCAGCAATCCCTCCTTCCTTCGGTTTC
 AACTGGACTTCTATCAGGTCTACTTCTTGGCCCTGGCAGCTGATTGGCTTCAGGCCCCCTAC
 CTCTATAAACTCTACCAGCATTACTACTTCTTGGAGGTCAAATTGCCATCCTCTATGTCTG
 TGGCCTTGCTCTACAGTCCTCTTTGGCCTAGTGGCCTCCTCCCTTGCGATTGGCTGGGTG
 GCAAGAATTCTTGCTGCTCTTCTCCCTGACTTACTCACTATGCTGCTTAACCAAACTCTCT
 CAAGACTACTTTGTGCTGCTAGTGGGGCGAGCACTTGGTGGGCTGTCCACAGCCCTGCTCTT
 CTCAGCCTTCGAGGCCTGGTATATCCATGAGCACGTGGAACGGCATGACTTCCCTGCTGAGT
 GGATCCAGCTACCTTTGCTCGAGCTGCCTTCTGGAACCATGTGCTGGCTGTAGTGGCAGGT
 GTGGCAGCTGAGGCTGTAGCCAGCTGGATAGGGCTGGGGCCTGTAGCGCCCTTTGTGGCTGC
 CATCCCTCTCCTGGCTCTGGCAGGGGCCCTTGGCCCTTCGAAACTGGGGGGAGAACTATGACC
 GGCAGCGTGCCTTCTCAAGGACCTGTGCTGGAGGCCTGCGCTGCCTCCTGTGCGACGCCGCGC
 GTGCTGCTGCTGGGCACCATAACAAGCTCTATTTGAGAGTGTATCTTTCATCTTTGTCTTCCT
 CTGGACACCTGTGCTGGACCCACACGGGGCCCCCTCTGGGCATTATCTTCTCCAGCTTCATGG
 CAGCCAGCCTGCTTGGCTCTTCCCTGTACCGTATCGCCACCTCCAAGAGGTACCACCTTCAG
 CCCATGCACCTGCTGTCCCTTGCTGTGCTCATCGTCGTCTTCTCTCTCTTCATGTTGACTTT
 CTCTACCAGCCCAGGCCAGGAGTCCGGTGGAGTCCTTCATAGCCTTTCTACTTATTGAGT
 TGGCTTGTGGATTATACTTTCCAGCATGAGCTTCCTACGGAGAAAGGTGATCCCTGAGACA
 GAGCAGGCTGGTGTACTCAACTGGTTCGGGTACCTCTGCACTCACTGGCTTGCCCTAGGGCT
 CCTTGTCCTCCATGACAGTGATCGAAAAACAGGCACTCGGAATATGTTGAGCATTTGCTCTG
 CTGTCATGGTGATGGCTCTGCTGGCAGTGGTGGGACTCTTACCGTGGTAAGGCATGATGCT
 GAGCTGCGGGTACCTTCACCTACTGAGGAGCCCTATGCCCTGAGCTG**TAA**CCCCACTCCAG
 GACAAGATAGCTGGGACAGACTCTTGAATTCAGCTATCCGGGATTGTACAGATCTCTCTGT
 GACTGACTTTGTGACTGTCTGTGGTTTCTCTGCCATTGCTTTGTGTTTGGGAGGACATGA
 TGGGGGTGATGGACTGGAAGAAGGTGCCAAAAGTCCCTCTGTGTTACTCCCATTTAGAAA
 ATAAACACTTTTAAATGATCAAAAAAAAAA

FIGURE 218

MLVTAYLAFVGLLASCLGLELSRCRAKPPGRACSNPSFLRFQLDFYQVYFLALAADWLQAPY
LYKLYQHYYFLEGQIAILYVCGLASTVLFGLVASSLVDWLGRKNSCVLFSLTYSLCCLTKLS
QDYFVLLVGRALGGLSTALLFSAFEAWYIHEHVERHDFPAEWIPATFARAAFWNHVLAVVAG
VAAEAVASWIGLGPVAPFVAAIPLLALAGALALRNWGENYDRQAFSRTCAGGLRCLLSDRR
VLLLGTIQALFESVIFIFVFLWTPVLDPHGAPLGIIFFSFMAASLLGSSLYRIATSKRYHLQ
PMHLLSLAVLIVVFSLFMLTFSTSPGQESPVESFIAFLLIELACGLYFPSMSFLRRKVIPET
EQAGVLNWFVRVPLHSLACLGLLVLHDSDRKTGTRNMFSCSAVMVMALLAVVGLFTVVRHDA
ELRVPSPTTEEPYAPEL

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 41-55, 75-94, 127-143, 191-213, 249-270, 278-299,
314-330, 343-359, 379-394, 410-430

FIGURE 219

GCGACGCGCGGCGGGGCGGGCGAGAGGAAACGCGGCGCGGGCCGGGCCCTGGAG**ATG**
 GTCCCCGGCGCCGCGGGCTGGTGTGTCTCGTGCTCTGGCTCCCCGCGTGCCTCGCGGCCA
 CGGCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCTTGGGGACATTCGATACA
 TCTTCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATGAGCAGATT
 CACCTTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACCTCAGCAACGGTTTCTTCATCCA
 GGACCAGATTGCTCTGGTGGAGAGGGGGGCTGCTCCTTCTCTCCAAGACTCGGGTGGTCC
 AGGAGCACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTAC
 GTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCTGCTCGG
 CCGAGACGGCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTT
 CCATCCCAGTCAATGTCAACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTC
 TGG**TAGA**AAGAGTTTGTCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCC
 AGGAATTTTGTACTTGGAAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGA
 AAAGGGTTTGGGCGTTGCTAGGCTGAAAGGGAAGCCACACCACTGGCCTTCCCTTCCCAGG
 GCCCCAAGGGTGTCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTA
 GAACCCGAACAAAAGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACT
 CACCTGGCTCCAGCCTCCCCTACCCAGGGTCTCTGCACAGTGACCTTACACAGCAGTTGTTGG
 AGTGGTTTAAAGAGCTGGTGTGTTGGGGACTCAATAAACCCCTCACTGACTTTTGTAGCAATAAA
 GCTTCTCATCAGGGTTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 220

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532
><subunit 1 of 1, 188 aa, 1 stop
><MW: 21042, pI: 5.36, NX(S/T): 2
MVPGAAGWCCLVLWLPACVAAHGFRIDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQ
IHLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDNAVDNDSF
YVEMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW
```

Signal peptide:

amino acids 1-20



FIGURE 221

TCTGCCTCCACTGCTCTGTGCTGGGATC**ATG**GAACTTGCACCTGCTGTGTGGGCTGGTGGTGA
TGGCTGGTGTGATTCCAATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTG
ACTGGGAAAATGCCCATCCTCTCCTACTGGCCCTACGGCTGTCACTGCGGACTAGGTGGCAG
AGGCCAACCCAAAGATGCCACGGACTGGTGTGCCAGACCCATGACTGCTGCTATGACCACC
TGAAGACCCAGGGGTGCGGCATCTACAAGGACAACAACAAAAGCAGCATACATTGTATGGAT
TTATCTCAACGCTATTGTTAATGGCTGTGTTAATGTGATCTATCTGGAAAATGAGGACTC
CGAA**TAA**AAAGCTATTACTAWTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

1-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
101-110
111-120
121-130
131-140
141-150
151-160
161-170
171-180
181-190
191-200
201-210
211-220
221-230
231-240
241-250

FIGURE 222

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76538
><subunit 1 of 1, 116 aa, 1 stop
><MW: 12910, pI: 6.41, NX(S/T): 1
MELALLCGLVVMAGVPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDW
CCQTHDCCYDHLKTQCGGIYKDNNKSSIHCMDSLQRYCLMAVFNVIIYLENEDSE
```

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 1-24

N-glycosylation site.

amino acids 86-89

N-myristoylation sites.

amino acids 20-25, 45-50

Phospholipase A2 histidine active site.

amino acids 63-70

FIGURE 223

CTCGCTTCTTCCTTCTGGATGGGGGGCCAGGGGGGCCAGGAGAGTATAAAGGCGATGTGGAG
GGTGCCCGGCACAACCAGACGCCCAGTCACAGGCGAGAGCCCTGGG**ATG**CACCGGCCAGAGG
CCATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTAT
GGCCCTGGAGGAGGCAAGTATTTCAGCACCACTGAAGACTACGACCATGAAATCACAGGGCT
GCGGGTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAAACTTGGAGACTCCTGGG
ACGTGAAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGCGAATAC
ATCACAAAAGTCTTTGTGCGCTTCCAAGCTTTCCTCCGGGGTATGGTCATGTACACCAGCAA
GGACCGCTATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCAGCCAAG
AGGGGCAGGTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGC
TTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCACCAAGTTAATCTCACATACTC
AGCAAACCTACCCGTGGGTGCGT**TAG**GGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGT
GTGGTGGTGGCTGATGGTACTGGAGTAACAGTTCGGGACGCTGAATCTGAATCCACCAATA
AATAAAGCTTCTGCAGAAAA

FIGURE 224

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19600, pI: 5.89, NX(S/T): 1
MHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKSVQVK
LGDSWDVKLGLALGGNTQEVTLQPGEYITKVFVAFQAFLRGMVMTSKDRYFYFGKLDGQISS
AYPSQEGQVLVGIYGQYQLLGIKISIGFEWNYPLEEPTTEPPVNLTYSANSPVGR
```

Signal peptide:

amino acids 1-22

100151451001

FIGURE 225

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCT
 GAACGGGTGCTCATACGGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAG
 CCCCAAATTGCGTGGAGAATAACATCATGTTTTCGATAAGAAAGAAATTGTAGGATCCAGTT
 TTTTTTTTAAACGCCCCCTCCCCACCCCCAAAAAACTGTAAGATGCAAAAACGTAATAT
 CCATGAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATT
 TATTTTGTCTTGGAGTGTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCT
 CCAAGGGGTCCAATTTTCTTCTGGGTGTCAGCGAGCCCTGACTCCTACAGTGACGCTG
 ACAGGGGCTGTCATGCAACTGGCCCCAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAA
 CAATACAAAGG**ATGG**GTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTT
 ATAGCCCCACTGTCTTACTGACAATGCTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTG
 TAGGTGTGAAGGCCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCTCAAGTA
 TATCTGCTGGTGTCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAACTTAAGTATAAT
 CAATTTAAAGGGCTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATAT
 TGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAAAGAA
 TCTCCTATTTTCTTAAACAATACCTTCAGACCTGTGACAAATTACGGAACCTGGATCTGTCC
 TATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTGGCGAAGCTGCTGAGTTT
 ACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACC
 TGGAACTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGC
 ATGATCAGACTCAAAGAACCTTACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCT
 TTTTCCAAGGTTGGTCAGCCTTCAGAACCCTTACTTGCACTGGAATAAAATCAGTGTCTATAG
 GACAGACCATGTCTGGACCTGGAGCTCCTTACAAGGCTTGATTTATCAGGCAATGAGATC
 GAAGCTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCGCGAATCTGCAGCGCTCAACCTGGA
 TTCCAACAAGCTCACATTTATTGGTCAAGAGATTTGGATTCTTGGATATCCCTCAATGACA
 TCAGTCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTGCTCCCTTGTAACCTGGCTG
 AAAAGTTTTAAAGGCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGG
 AGTAAATGTGATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGT
 TTGATCTGGCCAGGGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCAT
 GAGAGCAAAACCCCTTTGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGC
 TGACGCCGAGCACATCTCTTTCCATAAAATCATCGCGGCAGCGTGGCGCTTTTCTGTCCG
 TGCTCGTCATCTGTGGTTATCTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAG
 CTGCGACAGCGCTCCCTCATGCGAAGGCACAGGAAAAAGAAAGACAGTCCCTAAAGCAAT
 GACTCCCGACCCAGGAATTTTATGTAGATTATAAACCCCAACACCGGAGACCGGAGA
 TGCTGCTGAATGGGACGGGACCCTGCACCTATAACAAATCGGGCTCCAGGGAGTGTGAGGTA
TGAACCATTGTGATAAAAGAGCTCTTAAAGCTGGGAAATAGTGGTGTCTTATTGAACCT
 TGGTGACTATCAAGGGAACCGGATGCCCCCTCCCTTCCCTCTCCTCTCCTCTTGGTGG
 CAAGATCCTTCTTGTGCGTTTTAGTGCAATTCATAACTGGTCATTTCTCTCATACATA
 ATCAACCCATTGAATTTAAATACCAATCAATGTGAAGCTTGAACCTCCGGTTAATATAA
 TACCTATTGTATAAGACCTTTACTGATTGCCATTAATGTGCAATTTGTTTTAAGATAAACT
 TCTTTCATAGGTAAAAAAAAAA

FIGURE 226

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77301
><subunit 1 of 1, 513 aa, 1 stop
><MW: 58266, pI: 9.84, NX(S/T): 4
MGFNVIRLLSGSAVALVIAPTIVLLTMLSSAERGCPKGCRCGKMGVYCESQKLQEIPSSISAG
CLGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYF
LNNTFRPVNTLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELL
DLGYNRIRSLARNVFAGMIRLKELHLEHNQFSKLNALFPRLVSLQNLYLQWNKISVIGQTM
SWTWSSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLTFIGQEILDWSISLNDISLA
GNIWECSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLA
RALPKPTFFPKPLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKI IAGSVALFLSVLVI
LLVIYVSWKRYPASMKQLQQRSLMRRHRKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLN
GTGPCTYNKSGSRECEV
```

Important features of the protein:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 227

AGTTCCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTT
 TAAAT**ATG**TCAAGATCCAGACTTTTCAGTGTACCTCAGCGATCTCAACGATAGGGATCTTG
 TGTTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCGAAGAAGATGAAATGTGTGT
 AAATTATAATGACCAACACCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTGG
 TGGCAGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGA
 ATTGATTCTCACAGGCGACCATGGCAGTTTTTGTCTGTTGGAGACTTGGACTCTATTTATGG
 GACAGAAGCAGCTGTGAGTCCAACCTGTTGGAATTCACCTTCAAACCTCAAACCCCTGACCTAT
 ATCCTGTTCTGCTCCATGTTTTGGCCCTTAGGCTCCCCACCTCCATATGAAGAAATTTGA
 AAAACAACCT**TGA**TTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCCA
 AAACATCAAAATTTAGGAATAGTTATTTTCAGTTGTTGGAATGTCCAGAGATCTATTTCATATA
 GTCTGAGGAAGGACAATTCGACAAAAGAATGGATGTTGGAATAATTTTGGTCATGGAGATG
 TTTAAATAGTAAAGTAGCAGGCTTTTGATGTGTCACTGCTGTATCATACTTTTATGCTACAC
 AACCAAATTAATGCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTT
 TCCATCACAATTTAGGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTC
 CTAGCATGGGGTCCATAAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACAT
 GTCCAGAACCAGAACCAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAA
 GAAAAGTTTGGAGTTGAAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCT
 CAGCAAAAACAAGAGGTTTTATGCCCAACCTGAAGAGGAAGAAATTTGTAGATAGAAGGTGA
 AGGAGATTGCTGAAGATATAGAGCACATATAATGCCAACCGGGGAGAAAAGAAAATTTCCC
 CTTTTACAGTAATGAATGTGGCCTCCATAGTCCATAGTGTCTCTGGAGCCTCAGGGCTTG
 GCATTTATTGCAGCATCATGCTAAGAACCCTTCGCATAGGTATCTGTTCCCATGAGGACTGC
 AGAAGTAGCAATGAGACATCTCAAGTGGCATTTTGGCAGTGGCCATCAGCAGGGGGACAGA
 CAAAAACATCCATCACAGATGACATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAAT
 AAACATCAATAGATATCTAAAAA

FIGURE 228

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

><subunit 1 of 1, 146 aa, 1 stop

><MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSIAISTIGILCLPLFQLVLSDLPCCEEDEMCVNYNDQHPNGWYIWILLLLVLVA
ALLCGAVVLCLQCWLRPRIDSHRRTMAVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYP
VPAPCFGLGSPPPYEEIVKTT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 52-70

AAGCGGAGTAAAAATCTCCACAAGCTGGGAACAACCTCGTCCCAACTCCCACCACCGGCGT
 TTCTCCAGCTCGATCTGGAGGCTGCTTCGCCAGGTGTGGGACGCAAGTGCAGCCCCGCTTATTA
 GCTCTCGCTGCTCGCCCGGCTCAGAAGCTCCGTGGCCGCGCGCAAGTGCAGAGAAGCC
 ACGGCGCAGCTCAGTTCTCTTCTACTTTGGGAGAGAGAGAAAGTCAGATGCCCTTTTAAACT
 CCCCTCTCAAACTCATCTCCGGTGACTGAGTTAATAGATGGATCAACCTTGTCTGAAG
 ATGAAGAATATACAATATGAGGATATTTTTCTCTTTTTTTTCAAGTCTGTGATTTGTGGC
 TTACCTCAAGTTACCAATTTTTCAGTCAAGTCTGTGTTGTTGCTTCTTCAGAAATGTTTTTTT
 CAATCTCAAGAAAAAATATGTCCAGAAATTTGAGTTTCTGTTGCTGTGTTGATTTGGACTATT
 TGGGGATTGTAGTTACTGCACTATCTTTTCAACAACCAAGACATCAAGCAGTGTCAAGTT
 ACGTGAGCAAAATACAGACTTAAGCAAAAAGATGTAAAGCTTCAGCAGAGGAAAAAAGA
 ACACAGTGGATGTCGAGAACGGTGCTTCTATGGCAGGATATGCGGATCTGAAAGAACAAAT
 GCTGTCTTCTGGATGACATTTTGCAACGATTGGTGAAGCTGGAGAACAAAGTTGACTATAT
 TGTGTGAATGCTCAGCAGGCAACACACCAATGGTACTAGTGGGAATTTGGTGCCAGTAA
 CCACAAATTAAGAGCAAGATGCTCGGGCAGTATCAGATAGCAGTGAATACTACCTTGTGC
 TGCTCCATCCACTGTGGATTATATCTATGGCAGAAAAGCTTTATAATGCTGGCTTAGCAC
 AGAGCAATCTTTACAATAAAAGCTACACATTTTCAAGGAGTATGCTGATTATGAGAAC
 TCTAATTTCTGTACATAAAAAATTTAAAGATTTATTTGTTGCTTTGAGCAAGTCTGTTCAATG
 CTGTACTATGTCTCTAAAGAGAAATTTGGTAAGTGGTTGATGTTGAAGCAGATAGGTGAGT
 TTGTATAAATCTTTTGTGTTTGAGATCAAGCTGAAATGAAACAGTGAAGAAACATGGAAT
 ATTTCTATAACACATTTATTTAAGTATATAACACGTTTTTTGGACAAGTGAAGAAATGTTTAA
 TCATTCTGTCAATTTGTTCTCAATAGATTAAGTGTAGACTACGGCTATTTGAAAAAATGTG
 CTTATGTGACTATATTTTGTATTTCAATATGAGCAGGAAAGGAATATAATGTTGAAAA
 TAATGTTTGAATCATGACCCAAAGAAATGTATTGTTGCATATCCTTCAGAATAACTGA
 AGGTTAATTTGTATATTTTAAAAATTTACACTATAAAGATATAATCTTGAATGGGTAG
 CAGCCACTGTCCATTACCTATCGTAAACATTTGGGCAATTTAATAACAGCATTAATAAGTGT
 GTAACTCTAATCTTACTATTTATGAAGAATAAAGATATTTTTATGATGAGAGTAAACAT
 AAGTATTCAATGATTTTTCACATACATGAATGTTCACTTTAAAGGTTTAATCTTTGAGTGCT
 ATGCTATCAGGAAAGCACATTTTTCACATTTTGGGTTAAATTTTATTATATTGGTC
 TAGAGGAAAGGACTTTGGAGAAATGAACCTCTGAGGACTTTAGCCAGGTGTATATAATAAA
 GGTACTTTTGTGCTGCATTAATGTCTGGAAGTGTTAACATATATATATAAGATATCT
 CTTTATGAATTTTGAATTTGTATAACAGATGCATTAGATATTCATTTTATATAATGGCCAC
 TTAATAATAAGAACATTTAAATATAAATATAAGATGAAGATTGACTATCTTTTCAGGAAAAAGCT
 GTATATAGACAGGGAACCCCTAATCTTGGGTAATTCAGTATAAAACAAATATACATTTTAT
 TTAATATTTCCCTTGTAGCAACTATTAATGGCCATGGTGCCCTATATTTCTATAGATTTTAT
 CTCATAGTAGAAGTGTGTTAGTGCAGCTAGCTTCTAGATTAGCATATAGAAATTTAGATAT
 TGTATTGTTGCTCATTATATAATGCTACCCATGTAGCAATAATTACAATATTTTATTAATA
 TAAATATGTTGAATATTTGTTTCATGAAGAAGCAGATTTCCAATCTCTCTCTCTCTCTGTA
 CTGTCACCTTTTATGTGAAGAAATTAATATATATGTCATTGCCATTGCCAGGT

FIGURE 230

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648
><subunit 1 of 1, 140 aa, 1 stop
><MW: 15668, pI: 10.14, NX(S/T): 5
MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE
ENKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNL
VPVTTNKRTNVSGSIR
```

Important features of the protein:**Signal peptide:**

amino acids 1-26

FIGURE 231

CGCGGCCGGGCCGCCGGGGTGAAGCTGCCGAGGCGGCTGTGGCGCAGGCTTCCAGCCCCAC
CATCCCGTGGCCCTGCTGCTGCTGCTGGCCGTGAGTGGGGCCCCAGACAACCCGGCCATGCT
 TCCCCGGGTGCCAATGCAGGTGGAGACCTTCGGCCTTTTCAGACGCTTAGCCTGACTCGG
 GTGGATTGTAGCGGCCTGGGCCCCACATCATGCCGTGCCCATCCCTCTGGACACAGCCCA
 CTTGGACCTGTCTCCAACCGGCTGGAGATGGTGAATGAGTCGGTGTGGCGGGGCCGGGGT
 ACACGACGTGTGGCTGGCTGGATCTCAGCCACAACCTGCTCACCAGCATCTCACCACATGCC
 TTCTCCCGCCTTCGCTAGCTGGAGTCGCTTGACCTCAGCCACAATGGCCTGACAGCCCTGCC
 AGCCGAGAGCTTCACCAGCTCACCCTGAGCGACGTGAACCTTAGCCACAACAGCTCCGGG
 AGGTCTCAGTGCTGCTGCCTCACGACGCACAGTCAGGGCCGGGCACTACACGTGGACCTCTCC
 CACAACCTCATTCACCGCCTCGTGCCCCACCCACGAGGGGCCGGCCTGCCTGCGCCCAACAT
 TCAGAGCCTGAACCTGGCCTGGAACCGGCTCCATGCGGTGCCCCAACCTCCGAGACTTGCCCC
 TGCGCTACCTGAGCCTGGATGGGAACCCCTCTAGCTGTCAATTGGTCCGGGTGCCTTCGCGGGG
 CTGGGAGGCCTTACACACCTGTCTGTGCCAGCCTGCAGAGGCTCCCTGAGCTGGCGCCAG
 TGGCTTCCGTGAGCTACCGGGCCTGCAGGTCTCTGGACCTGTGCGGGCAACCCCAAGCTTAAC
 TGGCAGGAGCTGAGGTGTTTTAGGCCTGAGCTCCCTGCAGGAGCTGGACCTTTCCGGGCACC
 AACCTGGTGCCCTGCGCTGAGGCGCTGCTCCTCCACCTCCCGGCACTGCAGAGCGTCAGCGT
 GGGCCAGGATGTGCGGTGCCGCGCCTGGTTCGGGAGGGGCACCTACGCCGAGGCGCTGGCT
 CCAGCCCCAAGGTGCCCTGCACTGCGTAGACACCCGGGAATCTGCTGCCAGGGGCCCCACC
 ATCTTGT**GA**CAAAATGGTGGGCCAGGGCCACATAACAGACTGTGTCTCTGGGCTGCCTCAG
 GTCCCCAGTAACCTTATGTTCAATGTGCCAACACCAAGTGGGAGGCCCGAGGCCTATGTGGCA
 GCGTCACCACAGGAGTTGTGGGCCTAGGAGAGGCTTTGGACCTGGGAGCCACACCTAGGAGC
 AAAGTCTCACCCCTTTGTCTACGTTGCTTCCCCAAACCATGAGCAGAGGGAATTCGATGCCA
 AACCAGACTCGGGTCCCTCTGCTTCCCTTCCCCACTTATCCCCAAGTGCCTTCCCTCAT
 GCCTGGGCGGCCTGACCCGCAATGGGCAGAGGGTGGGTGGGACCCCTGCTGCAGGGCAGA
 GTTCAGGTCCACTGGGCTGAGTGTCCCTTGGGCCCCATGGCCCCAGTCACTCAGGGGGCAGTT
 TCTTTTCTAACATAGCCCTTTCTTTGCCATGAGGCCATGAGGCCCCGCTTCATCCTTTTCTAT
 TTCCCTAGAACTTAATGGTAGAAGGAATTGCAAGAATCAAGTCCACCCTTCTCATGTGAC
 AGATGGGAAACTGAGGCCTTGAGAAGGAAAAAGGCTAATCTAAGTTCCTGCGGGCAGTGGC
 ATGACTGGAGCACAGCCTCTCTGCCTCCAGCCCGGACCCAAATGCACCTTTCTGTCTCTCTA
 ATAAGCCCCACCCTCCCCGCTGGGCTCCCTTGTCTGCCCTTGCCTGTTCCCCATTAGCACA
 GGAGTAGCAGCAGCAGGACAGGCAAGAGCCTCACAAGTGGGACTCTGGGCCTCTGACCAGCT
 GTGCGGCATGGGCTAAGTCACTCTGCCCTTCGGAGCCTCTGGAAGCTTGGGACATCATGGTT
 CCAGCCTAGCCAGTTTCTACCCCTGGGTTGGGGTCCCCCAGCATCCAGACTGGAACCTACC
 CATTTTCCCTTGAGCATCTCTAGATGCTGCCCCAAGGAGTGTGTCAGTTCTGGAGCCTCA
 TCTGGCTGGGATCTCCAAGGGGCCCTCTTGATTGATGCTCCCACTGGCCCTGAGCAGCAGCAGC
 CCTTCTTACCTTCCAGGAATGCCGTGAAAGGAGACAAGGTCTGCCGACCCATGTCTATGC
 TCTACCCCCAGGGCAGCATCTCAGCTTCCGAACCTGGGCTGTTTCTTAGTCTTCAATTTA
 TAAAAGTTGTTGCCCTTTTAAACGGAGTGTCACTTCAACCGGCCCTCCCTACCCCTGCTGGC
 CGGGGATGGAGACATGTCAATTTGTAAGAGCAGAAAAAGGTTGCATTTGTTCACTTTTGTAA
 ATTGCTCTGGGCTGTGTTGGGCTGTTGGGGAAAGCTGGGCATCATGGCCACATGGGCATC
 ATGGGCTGGCCCCACAGAGACCCACAGGGCAGTGAGCTCTGTCTTCCCCACCTGCCTAGC
 CCATCATCTATCTAACCGGCTCTGATTTAATAAACACTATAAAAGGTTTAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 232

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77652
><subunit 1 of 1, 353 aa, 1 stop
><MW: 37847, pI: 6.80, NX(S/T): 2
MPWPLLLLLAVSGAQTRPCFPGCQCEVETFGLFDSFSLTRVDCSGLGPHIMPVPIPLDTAH
LDLSSNRLEMVNESVLAGPGYTTLAGLDLSHNLLTSISPTAFSRLRYLESLDLSHNGLTALP
AESFTSSPLSDVNLSHNQLREVSVSAFTTHSQGRALHVDLSHNLIHRLVPHPTRAGLPAPTI
QSLNLAWNRLHAVPNLRDLPLRYLSLDGNPLAVIGPGAFAGLGGLTHLSLASLQRLPELAPS
GFRELPGQLQVLDLSGNPKLNWAGAEVFSGLSSLQELDLSGTNLVPLPEALLHLPALQSVSV
GQDVRCRRLVRECTYPRRPGSSPKVPLHCVDTRESAARGPTIL
```

Signal peptide:

amino acids 1-16

Transmembrane domains:

amino acids 215-232, 287-304

FIGURE 233

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GATGGCGCAGCCACAGCTTCTGTGAGATTCGATTTCTCCCCAGTTCCTGTGGGTCTGAGG
 GGACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCA
 AAACAAGTTTTGACATTTCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCT
 GTTCCAGGCCCTTACCTGCTGGGCTAAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGC
 CACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTTTCAC
 GGGAGGCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAAGCCCTAAGATGAAAGCC
 TCTAGTCTTGCCTTCAGCCTTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGG
 ACTGAAGACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATG
 GATTTTCTGAGATACGGGGCAGTGTGCAAGCCAAAGATGGAACATTGACATCAGAATCTTA
 AGGAGGACTGAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCTCCTGCGCCATTT
 GCTAAGACTCTATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCC
 GGAAGATCAGCAGCCTCGCCAATTCTTTCTTACCATCAAGAAGACCTCCGGCTCTCTCAT
 GCCCACATGACATGCCATTGTGGGAGGAAGCAATGAAGAAATACAGCCAGATTTCTGAGTCA
 CTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAAGTACGACATTTCTT
 TGCAATGGATGGAGGAGACAGAATTAGGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGT
 CAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACT
 GTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGAT
 TGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGATAAGATTTTTGTAAATATCTT
 TCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTGTCTATTTAATGTATTT
 ATTTTTTTTACTTGGACATGAAACTTTAAAAAAATTCACAGATTATTTTATAACCTGACTAG
 AGCAGGTGATGTATTTTATACAGTAAAAAATAACCTTGTAATTTCTAGAAGAGTGGCT
 AGGGGGGTATTTCATTTGTATTCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGA
 TATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATT
 GCACATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTAT
 CTTCAGCCAGGAATCCTACAGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATA
 CCAAAAAAAAAAAAAAAAAA

FIGURE 234

```

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500
><subunit 1 of 1, 261 aa, 1 stop
><MW: 29667, pI: 8.76, NX(S/T): 0
MRQFPKTSFDISPEMSFSIYSLQVPVPG LTCWALTAEPGWGQNKGATTCATNSHSDSELRP
EIFSSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNL
QEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLDRVFKNYQTP
DHYTLRKISSLANSFLT IKKDLRLSHAHMTCHCGEEAMKKK SQILSHFEKLEPQAAVVKALG
ELDILLQWMEETE

```

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 235

CCGTTATCGTCTTGCGCTACTGCTGA**ATGT**CCGTCGCCGAGGAGGAGGAGAGGCTTTTGCCG
 CTGACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGGGCTACCGTGGC
 CGAGCTAGCAACCTTTCCCTGGATCTCACAAAACTCGACTCCAATGCAAGGAGAAGCAG
 CTCTTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCAC
 GCCCTAGGGATCATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCAT
 TTACAGACACGTAGTGTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTG
 TGTTTGGCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATG
 GCTGGTGTTATTGGCCAGTTTTTAGCCAATCCAACCTGACCTAGTGAAGGTTTCAGATGCAAAT
 GGAAGGAAAAAGGAACTGGAAGGAAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTG
 CAAAAATCTTAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAA
 AGAGCAGCACTGGTGAATATGGGAGATTTAACCCTTATGATACAGTGAACACTACTTGGT
 ATTGAATACACCACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGAC
 TGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCA
 CGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTGAGGCTGT
 TCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACCATCTTGGCTGAGAATGACCC
 CTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGATGAGTGGAGTCAGTCCA
 TTT**TAA**

FIGURE 236

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77568
><subunit 1 of 1, 323 aa, 1 stop
><MW: 36064, pI: 9.33, NX(S/T): 1
MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGA
RESAPYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFVGKSEDEH
YPLWKSIVIGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGGKPLRFRGVHHAFAKILAEGGI
RGLWAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTP
ADVIKSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLT
YEKIREMSGVSPF
```

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 237

CGGACGCGTGGGCGGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGC
 GCCTGAAGTCGGCGTGGGCGTTTGAAGAGCTGGGATACAGCATTAAATGAAAAATTTATGC
 TTAAGAAGTAAAAATGGCAGGCTTCCTAGATAATTTTCGTTGGCCAGAATGTGAATGTATTG
 ACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTGCGAGGTATATTGTTTTTTACAGGC
 TGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGAACCATGCCTT
 TCACACATGTGGTGATTTTCCACATGGCTTTCCTCATGATAAATGCTGTATCCAATGCTC
 AGGTGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTT
 TTCATTGGTTTTCATGTTGATGTTTGGGTCACTTATTGCTTCCATGTGGATTCTTTTTGGTGC
 ATATGTTACCCAAAATACTGATGTTTATCCGGGACTAGCTGTGTTTTTCAAATGCACCTTA
 TATTTTTTAGCACTCTGATCTACAAATTTGGAAGAACCAGAGCTATGGACCCTGAGATCAC
 TTCTTAAGTCACATTTTCTTTTGTATATTCTGTTTGTAGATAGGTTTTTTATCTCTCAGT
 ACACATTGCCAAATGGAGTAGATTGTACATTAATGTTTTGTTTCTTTACATTTTTATGTTT
 TGAGTTTGAAGTAGTTTTATGAAATTTCTTTATTTTCATTGCATAGACTGTTAATATGTA
 TATAATACAAGACTATATGAATTGGATAATGAGTATCAGTTTTTTATTCCTGAGATTAGAA
 CTTGATCTACTCCCTGAGCCAGGGTTACATCATCTTGTCATTTTAGAAGTAACCACTCTGT
 CTCTCTGGCTGGGCGACGGTGGCTCATGCCTGTAATCCAGCAGCTTTGGGAGGCCGAGGCGGG
 CCGATTGCTTGAGGTCAAGTGTTTGAGACCAGCCTGGCCAACATGGCGAAACCCATCTACT
 AAAAAATACAAAATTAGCCAGGCATGGTGGTGGTGCCTGTAATCCAGCTACCTGGGAGGC
 TGAGGCGAGGAGAATCGCTGAACCCGGGGGCGAGAGTTGCAGTGAGCTGAGTTTGCGCCAC
 TGCACCTCTAGCCTGGGGGAGAAAGTAAACTCCCTCTCAAAAAAAGACCACTCTCAGTATC
 TCTGATTTCTGAAGATGTACAAAAAATATAGCTTCATATATCTGGAATGAGCACTGAGCCA
 TAAAAGGTTTTTCAAGATTGTAACCTATTTTGGCCTAAAAATGAGGTTTTTTTTGGTAAAGA
 AAAAAATTTGTTCTTATGTATTGAAGAAGTGACTTTTATATAATGATTTTTTAAATGCC
 AAAGGACTAGTTTGAAAGCTCTTTTAAAAAGAAATCCTCTAATATGACTTTTATGTGAGAA

FIGURE 238

MAGFLDNFRWPECECIDWSERRNAVASVVAGILFFTGWWMIDAADVVPKPEQLNHAFHTCG
 VFSTLAFFMINAVSNAQVRGDSYESGCLGRTGARVWLFIGFMLMFGSLIASMWILFGAYVTQ
 NTDVYPGLAVFFQNALIFFSTLIYKFGRTTELWT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 23-42 (type II), 60-80, 97-117, 128-148

FIGURE 239

GTTGATGGCAAACCTCCTCAAAGGAGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCCA
 CTGGCGGCCCGCAACACTCCGTCTCACCCCTCTGGGCCCACTGCATCTAGAGGAGGGCCGTCT
 GTGAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTGAGAAGCTGGCCAGGGTGGT
 GGTCAAGTGGGTGAGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAG
 GGAAGTGGGAGCCTCGAGCCCTCGGGTGAAGCTGACCCCAAGCCACCTTCACCTGGACAG
GATGAGAGTGTGAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGA
 TGTATTATCGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCA
 GCCTCGCCACCAAGGAGATCCAGGTTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCC
 AGCCAACTACTTTGCGCTTTAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGT
 GCTTTGAAGACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATC
 GCCCTGGTGAATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGG
 AGATGTTATGCACCTAGTGAAATTCCTTAAAGAAATTCGGGGGGTGCCTGGTGTGGTGG
 CCTCCTACGACGATCCAGGGACCAAAATGAACGATGAAAGCAGGAACTCTTCTCTGACTTG
 GGGAGTTCCTACGCAAAACAACTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGA
 CCTCAGGGGTAAAAGCCCTTTGAGCAGTTCCTTAAAGAACAGCCAGACACAACAAATACG
 AGGGATGGCCAGAGCTGCTGGAGATGGAGGGCTGCATGCCCCGAAGCCATTTAGGGTGGC
 TGTGGCTCTTCCTCAGCCAGGGCCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCG
 GCAGGGGCTGAGGAGGAGGAGCAGGGGTGCTGCGTGAAGGTGCTGCAGGTCTTGCACGC
 TGTGTCGCGCCTCTCCTCCTCGGAAACAGAACCCTCCACAGCACATCTACCCGGAAGACC
 AGCCTCAGAGGGTCTTCTGGAACCAGCTGTCTGTGGAGAGAATGGGGTGTCTTCGTCAGGG
 ACTGTGACGGCTGGTCTGAGGAAGGACAACTGCCAGACTTGAGCCCAATTAAATTTTA
 TTTTGTCTGGTTTTGAAAAAAAAAAAAAAAAAAAA



FIGURE 240

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59814
<subunit 1 of 1, 224 aa, 1 stop
<MW: 24963, pI: 9.64, NX(S/T): 1
MRVSGVLRLLALIFAIVTTWMFIRSYMSFSMKTIRLPRWLAASPTKEIQVKYKCGLIKPCP
ANYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSG
DVMHLVKFLKEIPGGALVLVASYDDPGTKMNDESRKLFSDLGSSYAKQLGFRDSWVFIGAKD
LRGKSPFEQFLKNSPDTNKEYEGWPELLEMEGCMPPKPF
```

Important features:**Signal peptide:**

amino acids 1-15

ATP/GTP-binding site motif A (P-loop).

amino acids 184-191

N-glycosylation site.

amino acids 107-110

FIGURE 241

101537-121101

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCTGGGGAT
 CCAGAAACCCATGATACCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACA
 GCAAGAGAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTC
 TCACTCCTCCCTCCCTCTCTCTCTGCTGTCTAGTCTCTAGTCTCAAATCCCAGTCCC
 CTGACCCCCCTTCTGGGACACT**ATG**TTGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCC
 TGGCTGCAGATGGGGGTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGGCCA
 GCCTCTTACCCTGAGTGTGGAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGT
 GACATTTGACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCAGC
 CTTTGGACCTGCACAACAATGGCCACACAGTGCAACTCTCTCTGCCCCTTACCCTGTATCTG
 GGTGGACTTCCCCGAAAATATGTAGCTGCCCAGCTCCACCTGCATGGGGTCAGAAAGGATC
 CCCAGGGGGGTGAGAACACCAGATCAACAGTGAAGCCACATTTCAGAGCTCCACATTGTAC
 ATTATGACTCTGATTCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCCTCAGGGCCTGGCT
 TCCCTGGGCATCCTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAG
 TCACITGTCATGAAGTCAGGCATAAAGATCAGAAGACCTCAGTGCCCTCCCTTCAACCTAAGAG
 AGCTGCTCCCCAACAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACTCCCCCT
 TGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAGATTTCAATGGAACAGCT
 GGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGA
 ACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTTCATCCAAGCAGGA
 TCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGCTGTCTCTG
 CCTTCTCCTGGCTGTTTTATTTCATTGCTAGAAAGATTCGGAAGAAGAGGCTGGAAAACCGAA
 AGAGTGTGGTCTTCACCTCAGCACAGCCACGACTGAGGCAT**TAA**ATTCTTCTCAGATACCA
 TGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAATGGGGTGTAGGATCTG
 GCCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAG
 GAATGGACCCAGGCTGTCAATTCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTA
 GGAGGAAATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAATCTGTTTAGTTGCAGGG
 GAAGTTTGGGATATACCCCAAAGTCTCTACCCCTCACTTTTATGGCCCTTTCCCTAGATA
 TACTGCGGATCTCTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTGTATCAATA
 TATTGGAAATTAAAGTTTCTGACTTT

FIGURE 242

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812
><subunit 1 of 1, 337 aa, 1 stop
><MW: 37668, pI: 6.27, NX(S/T): 1
MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWFPASYPECGNNAQSPIDIQTDSVTFDPDLP
ALQPHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPPGGSEHQ
INSEATFAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRH
KDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPCYQSVLWTVFYRRSQISMEQLEKLQGTLL
FSTEEEPSKLLVQNYRALQPLNQRMVFASFIIQAGSSYTTGEMLSLGVGILVGCCLLLAVYF
IARKIRKKRLENRKSVVFTSAQATTEA
```

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

10015371.123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

AATTTTTCACCAGAGIAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTC
GTGGACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGG
ATCAACTCGGTTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTC
CGGATCAGGGAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCTCTCTTTAAGTCTG
ATACCATTAAACACAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGG
AATGACACCTGGTACCCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAAC
TGCACCCACATGTGTTACCAATTTTTGTACACAACCTTGAGAGCCAGGGCACTATCCTAAGC
TCAGAGGAATTGCCACAAATCTTCACGAGCCTCATCATCCATTCCCTGTGTTCCCGGAGGCAT
CCTGCCCACCAGTCAGGCAGGGGCTAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAG
GAGCAGGTGTAAATCCTGCCACCCAGGGAACCCAGCAGGCCGCTCCCAACTCCCAGTGGC
ACAGATGACGACTTGTGAGTGACCACCCCTGCAGGCATCCAAGGAGCACACATGCCATCGA
GGAAGCCACCACAGAATCAGCAAATGGAATTTCAGTAAGCTGTTTCAAATTTTTTCAACTAAG
CTGCCTCGAATTTGGTGATACATGTGAATCTTTATCATTGATTATATTATGGAATAGATTGA
GACACATTGGATAGTCTTAGAAGAAATTAATCTCTAATTTACCTGAAAATATTCTTGAATTT
TCAGAAAAATTGTTCTATGTAGAGAAATCCCACTTTTAAAAACAATAATTCAATGGATAAAT
CTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATATTAACACATATTGAAAA
ACTGGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA

FIGURE 244

MRSTILLFCLLGSTRSLPQLKPAIGLPPTKLPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDLLHLLNPAAGMTPGTQTHPLTLGGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLEPAGGAGVNPATQGTTPAGRLLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 245

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCG
 GAGCCAGACGCTGACCACGTTCTCTCTCGGTCTCCTCCGCTCCAGCTCCGCGCTGCCCCG
 GCAGCCGGGAGCCCATGCAGCCCCAGGGCCCCGCCCTCCCCGAGCGGCTCCGCGGCCTCC
 TGCTGCTCTGCTGCTGCAGCTGCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAG
 CAAAAGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGG
 GCCAGCAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTG
 GGATCCCAGGTGCGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGAAAGCTTTGAG
 GAGTCCTGGACACCCAACTACAAGCAGTGTTCATGGAGTTCATTGAATTATGGCATAGATCT
 TGGGAAAATTGCGGAGTGTACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGT
 TCAGTGGCTCACTTCGGCTAAAATGCAGAAATGCATGCTGTGAGCGTTGGTATTTTCACATTC
 AATGGAGCTGAATGTTCAGGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAG
 CCTGAAATGAATTCACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAG
 GAATTGGTGTGATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTCAGATTACCCAAAA
 GGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAATTA
AATGCTTTAATTTTCATTTGCTACCTCTTTTTTTATTATGCCTTGGAATGGTTCACTTAAAT
 GACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGA
 CCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTTCATTTTGCTTCAATCAAAGT
 GGTTCATATATTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTA
 TAATTTGGAATATTGTTGGTCTTTTGTTTTTCTCTTAGTATAGCATTTTTTAAAAAATA
 TAAAAGCTACCAATCTTTGTACAATTTGTAATGTTAAGAATTTTTTTTATATCTGTTAAAT
 AAAAATTATTTCCAACA

FIGURE 246

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393
><subunit 1 of 1, 243 aa, 1 stop
><MW: 26266, pI: 8.43, NX(S/T): 1
MRPQGPAAASPQRLRGLLLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQGPGAGV
PGRDGSFGANVIPGTPGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWSVSRIIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217